

mg

78224

184291

STIC-Biotech/ChemLib

From: Seharaseyon, Jegatheesan  
Sent: Wednesday, April 05, 2006 11:25 AM  
To: STIC-Biotech/ChemLib  
Subject: Re:10/067832

Importance: High

Please search SEQ ID NO: 2 and protein encoded by SEQ ID NO 1 and 13.

Thanks.

J.Seharaseyon  
Art Unit 1647  
Remsen 4C61  
Mailbox 4C70  
Phone: (571)-272-0892  
Fax: (571)-273-0892

RECEIVED  
APR - 5 2006  
STIC-BIOTECH/ChemLib  
(STIC)

\*\*\*\*\*  
Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIS: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using **frame\_plus.n2p model**  
Run on: April 7, 2006, 23:31:58 ; Search time 31.8263 Seconds  
(without alignments)  
1021.610 Million cell updates/sec

Title: US-10-067-832D-1  
Perfect score: 665  
Sequence: 1 gttccaggagattacagctc.....tttgatgctgcaaaaaaaa 370

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat\_07042006\_175609\_9682/app\_query.fasta.1  
-DB=A Geneseq -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p  
-USER=US10067832 @CGN 1.1 476 @runat\_07042006\_175609\_9682 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_21.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*  
9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	62.9	89	4 AAM41466	Aam41466 Human pol
2	418	62.9	115	5 ABP41331	Abp41331 Human ova
3	394	59.2	73	2 AAY08413	Aay08413 P. Obesus
4	394	59.2	73	3 AAB36290	Aab36290 Israeli s
5	394	59.2	73	4 AAM39680	Aam39680 Human pol
6	339	51.0	73	4 ABB59910	Abb59910 Drosophil
7	313	47.1	73	3 AAG41925	Aag41925 Arabidops
8	313	47.1	96	3 AAG41924	Aag41924 Arabidops
9	312	46.9	73	3 AAG40885	Aag40885 Zea mays

10	300	45.1	73	3 AAG27281	Aag27281 Zea mays
11	285	42.9	73	8 ADT58384	Adt58384 Plant pol
12	252	37.9	89	5 ADH32312	Adh32312 Novel Yea
13	252	37.9	320	6 ABR83620	Abr83620 HUB1-GFP
14	163	24.5	33	2 AAY08414	Aay08414 Human bea
15	163	24.5	33	8 ADR43346	Adr43346 Human bea
16	160.5	24.1	112	8 ADX97089	Adx97089 Plant ful
17	160.5	24.1	112	8 ADX78384	Adx78384 Plant ful
18	158	23.8	112	8 ADY06931	Ady06931 Plant ful
19	158	23.8	113	8 ADY05953	Ady05953 Plant ful
20	158	23.8	128	8 ADX96955	Adx96955 Plant ful
21	147.5	22.2	32	3 AAB36291	Aab36291 Human bea
22	138	20.8	27	4 AAM21187	Aam21187 Peptide #
23	138	20.8	27	4 ABB43508	Abb43508 Peptide #
24	138	20.8	27	4 AAM37402	Aam37402 Peptide #
25	138	20.8	27	4 ABB26468	Abb26468 Protein #
26	138	20.8	27	4 AAM77251	Aam77251 Human bon
27	138	20.8	27	4 AAM64443	Aam64443 Human bra
28	138	20.8	27	5 ABG46287	Abg46287 Human pep
29	138	20.8	110	8 ADY04784	Ady04784 Plant ful
30	138	20.8	124	8 ADY04557	Ady04557 Plant ful
31	137	20.6	40	9 ADY95411	Ady95411 Yeast ubi
32	137	20.6	104	5 ADH32742	Adh32742 Yeast smo
33	133	20.0	40	9 ADY95410	Ady95410 Yeast ubi
34	120	18.0	32	9 ADV95409	Adv95409 Yeast ubi
35	115	17.3	32	9 ADX95412	Adx95412 Yeast ubi
36	88.5	13.3	854	9 ADX39709	Adx39709 HIV Env p
37	85.5	12.9	857	9 ADX39713	Adx39713 HIV Env p
38	85	12.8	857	9 ADX39712	Adx39712 HIV Env p
39	83.5	12.6	638	9 ADZ07768	Adz07768 HIV CON-A
40	83.5	12.6	854	9 ADZ07767	Adz07767 HIV CON-A
41	83.5	12.6	854	9 ADZ07813	Adz07813 HIV CON-A
42	82.5	12.4	853	9 ADZ04172	Adz04172 HIV prote
43	82	12.3	831	9 ADZ07806	Adz07806 HIV CON-F
44	80	12.0	123	8 ADY05551	Ady05551 Plant ful
45	80	12.0	125	8 ADY05562	Ady05562 Plant ful

ALIGNMENTS

RESULT 1  
AAM41466  
ID AAM41466 standard; protein; 89 AA.  
XX  
AC AAM41466;  
XX  
AC  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6397.  
XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX

XX  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR

```

PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60622.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 6397; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 89 AA;
SQ
Alignment Scores:
Pred. No.: 1,03e-41 Length: 89
Score: 418.00 Matches: 78
Percent Similarity: 96.3% Conservative: 1
Best Local Similarity: 95.1% Mismatches: 3
Query Match: 62.9% Indels: 0
DB: Gaps: 0
US-10-067-832D-1 (1-370) x AAM41466 (1-89)
QY 2 TTCACGAGATTACACTCCAGCCACATGATTGAGTGGTTTGCACACCGCTTAGCA 61
Db 8 PheGluAlaIleArgAlaProAlaArgMetIleGluValValCysAenAspArgLeuGly 27
QY 62 AAGAAGTCGGCTTAAGTGCACACCGATGACCCATCGGGACTTGAAGAACTGATA 121
Db 28 LysLysValArgValLysCysAenThrAspThrIleGlyAspLeuLysLeuIle 47
QY 122 GCGGCCCAACTGGCACTCGTTGGAATAGATCGTTCTTAAAGTGGTACACGATTTT 181
Db 48 AlaAlaGlnThrGlyThrArgIlePheLysIleValLeuLysLysIleThrIlePhe 67
QY 182 AAGGACCATGTATCTCTGGAGATTATGAATCCACATGGGATGACCTGGAGCTTTAT 241
Db 68 LysAspHisValSerLeuGlyAspTyrGluIleHisAspGlyMetAenLeuGluLeuTyr 87
QY 242 TACCAG 247
Db 88 TyrGln 89
RESULT 2
ABP41331
ID ABP41331 standard; protein; 115 AA.
XX
XX ABP41331;
XX
XX 22-AUG-2002 (first entry)
XX
XX Human ovarian antigen HACNC39, SEQ ID NO:2463.
XX

```

---

```

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX
XX Homo sapiens.
XX
XX WO200200677-A1.
XX
XX 03-JAN-2002.
XX
XX 07-JUN-2001; 2001WO-US018569.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54408.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX
XX Claim 11; SEQ ID NO 2463; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention. Note: The
XX sequence data for this patent did not form part of the printed from WIPO
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 115 AA;
XX
Alignment Scores:
Pred. No.: 1.1e-41 Length: 115
Score: 418.00 Matches: 78
Percent Similarity: 96.3% Conservative: 1
Best Local Similarity: 95.1% Mismatches: 3
XX

```



Query Match: 62.9% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-1 (1-370) x ABP41331 (1-115)

QY 2 TTCAGGAGATTACAGCTCCAGCCACCAATGATTGAGTGGTTTGCACGACCGTCTAGGA 61  
| | | | |  
DB 34 PheGluAlaIleArgAlaProAlaArgMetIleGluValValCysAsnAspArgLeuGly 53  
| | | | |  
QY 62 AAGAAAGTCGCGTTAAGTGCACACCGATGACACCATCGGGGACTTGAAGAACTGATA 121  
| | | | |  
DB 54 LysLysValArgValLysCysAsnThrAspAspThrIleGlyAspLeuLysLysLeuIle 73  
| | | | |  
QY 122 GCGGCCCAACTGCACTCGTTCGTAAGTAAAGTCTTAAAGTGGTACACGATTTT 181  
| | | | |  
DB 74 AlaAlaGlnThrGlyThrArgTrpAsnLysIleValLeuLysLysTrpThrIlePhe 93  
| | | | |  
QY 182 AAGGACCATGATCTCTCGGAGATTATGAATCCACGATCGGATGAACCTGGAGCTTTAT 241  
| | | | |  
DB 94 LysAspHisValSerLeuGlyAspTrpGluIleHisAspGlyMetAsnLeuGluLeuTrp 113  
| | | | |  
QY 242 TACCAG 247  
| | | | |  
DB 114 TyrGln 115  
| | | | |

RESULT 3  
AAY08413  
ID AAY08413 standard; protein; 73 AA.  
AC AAY08413;  
XX  
DT 24-JUL-1999 (first entry)  
XX  
DE P. obesus beacon protein.  
XX  
KW Beacon; hypothalamus; obese; lean; agonist; antagonist; treatment;  
KW obesity; anorexia; weight maintenance; energy imbalance; diabetes;  
KW metabolic syndrome; dyslipidemia; hypertension; insulin resistance;  
KW medicament; livestock; diagnosis.  
XX  
OS Peannomys obesus.  
XX  
PN WO9923217-A1.  
XX  
PD 14-MAY-1999.  
XX  
PF 30-OCT-1998; 98WO-AU000902.  
XX  
PR 31-OCT-1997; 97AU-00000117.  
XX  
PR 11-NOV-1997; 97AU-00000323.  
XX  
XX (ITDI-) INT DIABETES INST.  
PA (UYDE-) UNIV DEAKIN.  
XX  
XX Zimmet PZ, Collier G;  
PI  
XX WPI; 1999-337484/28.  
DR N-PSDB; AAX57359.  
XX  
XX New gene encoding a beacon protein associated with modulation of obesity,  
PT diabetes and metabolic energy levels.  
XX  
PS Claim 2; Page 50; 85pp; English.  
XX  
CC This invention describes a novel beacon protein and its encoding nucleic  
CC acid which is expressed in larger amounts in hypothalamus tissue of obese  
CC animals compared to lean animals. Agonists and antagonists of beacon can  
CC be used to treat obesity, anorexia, weight maintenance, energy imbalance,  
CC diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin  
CC resistance. The beacon protein, itself is used to manufacture medicaments  
CC for treatment of obesity, anorexia, energy imbalance or diabetes. The  
CC treatment is contemplated for both human and animals, such as those  
CC important to the livestock industry. The antibody and polynucleotides are

CC useful in diagnosis of conditions as above  
XX  
SQ Sequence 73 AA;  
Alignment Scores:  
Pred. No.: 8,258-39 Length: 73  
Score: 394.00 Matches: 73  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 59.2% Indels: 0  
DB: Gaps: 0

US-10-067-832D-1 (1-370) x AAY08413 (1-73)

QY 29 ATGATTGAGTGGTTTTCACACGACCGTCTAGGAAAGAAAGTCCGCTTAAGTGCACACC 88  
| | | | |  
DB 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
| | | | |  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCATCTGTTGAAT 148  
| | | | |  
DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40  
| | | | |  
QY 149 AAGATCGTCTTAAAGTGGTACACGATTTTAAAGGACCATGATCTCTGGGAGATTAT 208  
| | | | |  
DB 41 LysIleValLeuLysLysTrpThrIlePheLysAspHisValSerLeuGlyAspTrp 60  
| | | | |  
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247  
| | | | |  
DB 61 GluIleHisAspGlyMetAsnLeuGluLeuTrpGln 73  
| | | | |

RESULT 4  
AAB36290  
ID AAB36290 standard; protein; 73 AA.  
XX  
AC AAB36290;  
XX  
DT 23-FEB-2001 (first entry)  
XX  
DE Israeli sand rat beacon ligand.  
XX  
KW Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;  
KW ligand.  
XX  
OS Peannomys obesus.  
XX  
PN WO200064931-A1.  
XX  
PD 02-NOV-2000.  
XX  
PF 19-APR-2000; 2000WO-AU000342.  
XX  
PR 23-APR-1999; 99AU-00009919.  
XX  
PR 24-MAR-2000; 2000AU-00006454.  
XX  
XX (AUTO-) AUTOGEN PTY LTD.  
PA  
XX Collier G, Walder K, Zimmet P;  
PI  
XX WPI; 2000-687311/67.  
DR N-PSDB; AAC81767.  
XX  
XX Ligand of beacon protein useful for treating obesity, anorexia, energy  
PT imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and  
PT insulin resistance.  
XX  
PS Claim 3; Fig 1; 67pp; English.  
XX  
CC The present invention is related to the isolation of a ligand known as  
CC beacon from the Israeli sand rat. Beacon is associated with the  
CC regulation of energy balance, and the protein, its coding sequence and  
CC analogues can be used in the treatment of diabetes, obesity, anorexia,  
CC energy imbalance, metabolic syndrome, dyslipidemia, hypertension and  
CC insulin resistance. In addition, they can be used in agriculture to

```

CC produce leaner animals
XX
SQ Sequence 73 AA;

Alignment Scores:
Pred. No.: 8,25e-39 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 3 Gaps: 0

US-10-067-832D-1 (1-370) x AAM36290 (1-73)
QY 29 ATGATTGAGTGGTTTGCACACCGCTTAGGAAAGAAAGTCCGGTTAAGTGCACACC 88
DB 1 MetIleGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGCACTCGTTGGAT 148
DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
DB 41 LysIleValLeuLysLysTrpThrIlePheLysAspHisValSerLeuGlyAspTyr 60

QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
DB 61 GluIleHisAspGlyMetAsnLeuGluLeuTyrGln 73

RESULT 5
AAM39680
ID AAM39680 standard; protein; 73 AA.
AC AAM39680;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2825.
XX
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSB-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
WPI; 2001-442253/47.
DR
N-PSDB; AAI58836.

```

```

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX
PS Example 4; SEQ ID NO 2825; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with neotropic.
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 73 AA;

Alignment Scores:
Pred. No.: 8,25e-39 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 4 Gaps: 0

US-10-067-832D-1 (1-370) x AAM39680 (1-73)
QY 29 ATGATTGAGTGGTTTGCACACCGCTTAGGAAAGAAAGTCCGGTTAAGTGCACACC 88
DB 1 MetIleGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGCACTCGTTGGAT 148
DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
DB 41 LysIleValLeuLysLysTrpThrIlePheLysAspHisValSerLeuGlyAspTyr 60
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
DB 61 GluIleHisAspGlyMetAsnLeuGluLeuTyrGln 73

RESULT 6
ABB59910
ID ABB59910 standard; protein; 73 AA.
XX
AC ABB59910;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 6522.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.

```

XX PA (PEKE ) PE CORP NY.  
XX FI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX DR N-PSDB; ABL04013.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from *Drosophila* and for elucidating cell signaling and cell-cell  
XX interactions.  
XX PS Disclosure; SEQ ID NO 6522; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from *Drosophila*. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 73 AA;  
  
XX Alignment Scores:  
XX Pred. No.: 4.2e-32 Length: 73  
XX Score: 339.00 Matches: 63  
XX Percent Similarity: 90.4% Conservative: 3  
XX Best Local Similarity: 86.3% Mismatches: 7  
XX Query Match: 51.0% Indels: 0  
XX DB: 4 Gaps: 0  
  
XX US-10-067-832D-1 (1-370) x ABB59910 (1-73)  
XX QY 29 ATGATTGAGTGGTGTTCGACAGCGCTCTAGGAAAGAAAGTCCGGTTAAGTCAACACC 88  
XX Db 1 MetileGlulileThrCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnPro 20  
XX QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAACTCGCACTCGTGGGAT 148  
XX Db 21 AspAspThrIleGlyAspLeuLysLysLeuLysLeuLysLeuLysLeuLysLeuLys 40  
XX QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATCTATCTCTGGGAGATTAT 208  
XX Db 41 LysileValLeuLysLysIleThrIlePheLysAspProileArgLeuSerAspTyr 60  
XX QY 209 GAATCCAGATCGGATGACCTGGAGCTTTATTACAG 247  
XX Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyTyTyGln 73  
  
XX RESULT 7  
XX AAG41925  
XX ID AAG41925 standard; protein; 73 AA.  
XX AC AAG41925;  
XX XX  
XX DT 18-OCT-2000 (first entry)  
XX XX  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52222.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX XX  
XX OS Arabidopsis thaliana.  
XX XX  
XX PN EP1033405-A2.  
XX XX  
XX PD 06-SEP-2000.  
XX XX

PF 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 23-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145113P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150366P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.

PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. No.: 6,23e-29 Length: 73  
Score: 313.00 Matches: 58  
Percent Similarity: 88.9% Conservative: 6  
Best Local Similarity: 80.6% Mismatches: 8  
Query Match: 47.1% Indels: 0  
DB: 3 Gaps: 0  
US-10-067-832D-1 (1-370) x AAG41925 (1-73)

QY 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAAGAAAGTCCGCTTAAGTGCAACACC 88  
|||||  
Db 1 MetIleGluValValLeuAenAspArgLeuGlyLysValArgValLysCysAenAsp 20  
QY 89 GATGACACCTCGGGACTTGAGAACTGATAGCGGCCCAACTGGCACTCGTTGGAT 148  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuValAlaGlnThrGlyThrArgAlaGlu 40  
QY 149 AAGATCGTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGAGATTAT 208  
|||||  
Db 41 LysIleArgIleGlnLysTrpYrAsnIleLysAspHisIleThrLeuLysAspTyr 60  
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTAC 244  
|||||  
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

RESULT 8  
AAG41924  
ID AAG41924 standard; protein; 96 AA.  
XX  
AC AAG41924;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52221.  
XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128334P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135253P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 18-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 26-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.

PR 18-AUG-1999; 99US-0149426P.

PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149722P.

PR 23-AUG-1999; 99US-0149929P.

PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.

PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.

PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153750P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. No.: 6,7e-29  
Score: 313.00  
Percent Similarity: 88.9%  
Best Local Similarity: 80.6%  
Query Match: 47.1%  
DB: 3

US-10-067-832D-1 (1-370) x AAG41924 (1-96)

QY 29 ATGATTGAGTGGTTTTCACACGACCGCTAGGAAAGAAAGTCCGCGTTAAGTGCACACC 88  
Db 24 MetileluValValLeuAenAspArgGlyValysValargVallysCysAenAsp 43  
QY 89 GATGACACCATCGGGACTTTGAAGAACTGATAGCGGCCCAAACTGCACCTCGTTGGAAT 148  
Db 44 AspAspThrileGlyAspLeuLysLeuValAlaAlaGlnThrGlyThrArgAlaGlu 63  
QY 149 AAGATCGTTCTTAAAGTGGTACACCATTTTAAAGCACCATGTATCTCTCGGAGATTAT 208  
Db 64 LysileArgileGlnLysTrpTyraenileTyrlsAspHisileThrLeuLysAspTy 83  
QY 209 GAAATCCACCATGGGTAACCTCGAGCTTTATTAC 244  
Db 84 GlulleHisaspGlyMetGlyLeuGlulueuTyrrTy 95

RESULT 9  
AAG40885  
ID AAG40885 standard; protein; 73 AA.  
XX AC AAG40885;  
XX DT 18-OCT-2000 (first entry)  
XX DE Zea mays protein fragment SEQ ID NO: 50790.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX OS Zea mays subsp. mays.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0128645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.

[illegible]

PR	20-AUG-1999,	99UB-01499299P,
PR	23-AUG-1999,	99UB-01499029P,
PR	23-AUG-1999,	99UB-01499030P,
PR	25-AUG-1999,	99UB-0150566P,
PR	26-AUG-1999,	99UB-0150884P,
PR	27-AUG-1999,	99UB-0151065P,
PR	27-AUG-1999,	99UB-0151066P,
PR	27-AUG-1999,	99UB-0151080P,
PR	30-AUG-1999,	99UB-0151303P,
PR	31-AUG-1999,	99UB-0151438P,
PR	01-SEP-1999,	99UB-0151930P,
PR	07-SEP-1999,	99UB-0152363P,
PR	10-SEP-1999,	99UB-0153070P,
PR	13-SEP-1999,	99UB-0153758P,
PR	15-SEP-1999,	99UB-0154018P,
PR	16-SEP-1999,	99UB-0154039P,
PR	20-SEP-1999,	99UB-0154303P,
PR	22-SEP-1999,	99UB-0154477P,
PR	23-SEP-1999,	99UB-0155133P,
PR	24-SEP-1999,	99UB-0155486P,
PR	28-SEP-1999,	99UB-0156559P,
PR	28-SEP-1999,	99UB-0156458P,
PR	29-SEP-1999,	99UB-0156596P,
PR	04-OCT-1999,	99UB-0157117P,
PR	06-OCT-1999,	99UB-0157753P,
PR	06-OCT-1999,	99UB-0157865P,
PR	07-OCT-1999,	99UB-0158029P,
PR	08-OCT-1999,	99UB-0158232P,
PR	12-OCT-1999,	99UB-0158369P,
PR	13-OCT-1999,	99UB-0159293P,
PR	13-OCT-1999,	99UB-0159294P,
PR	13-OCT-1999,	99UB-0159295P,
PR	14-OCT-1999,	99UB-0159329P,
PR	14-OCT-1999,	99UB-0159333P,
PR	14-OCT-1999,	99UB-0159333P,
PR	14-OCT-1999,	99UB-0159637P,
PR	14-OCT-1999,	99UB-0159638P,
PR	18-OCT-1999,	99UB-0159584P,
PR	21-OCT-1999,	99UB-0160741P,
PR	21-OCT-1999,	99UB-0160767P,
PR	21-OCT-1999,	99UB-0160768P,
PR	21-OCT-1999,	99UB-0160770P,
PR	21-OCT-1999,	99UB-0160814P,
PR	22-OCT-1999,	99UB-0160815P,
PR	22-OCT-1999,	99UB-0160980P,
PR	22-OCT-1999,	99UB-0160981P,
PR	22-OCT-1999,	99UB-0160989P,
PR	25-OCT-1999,	99UB-0161404P,
PR	25-OCT-1999,	99UB-0161405P,
PR	25-OCT-1999,	99UB-0161406P,
PR	26-OCT-1999,	99UB-0161359P,
PR	26-OCT-1999,	99UB-0161360P,
PR	26-OCT-1999,	99UB-0161361P,
PR	28-OCT-1999,	99UB-0161920P,
PR	28-OCT-1999,	99UB-0161992P,
PR	29-OCT-1999,	99UB-0161993P,
PR	29-OCT-1999,	99UB-0162142P,

**Alignment Scores:**

Assignment Scores:	8.24e-29
Pred. No.:	312.00
Score:	88.9%
Percent Similarity:	80.6%
Best Local Similarity:	46.9%
Query Match:	3
DB:	

Length:	73
Matches:	58
Conservative:	6
Mismatches:	8
Indels:	0
Gaps:	0

US-10-067-832D-1 (1-370) x AAG40885 (1-73)

Qy 29 ATGATTGAGTGGTTTGCAACGACCCTCTAGAGAAAGAAGTCGCCGTTAAGTGCAACACC 88  
|||  
Db 1 MetIleGluValValIleuaenaphargLeuGLYLYsLyvaLaTgVa.LysCyaaenGlu 20  
|||  
Qy 89 GATGACACCATCGGGGACTTGAGAAACTGATAGCGGCCCAACTGGCACCTCGTTGGGAAT 141  
|||





PR 10-AUG-1999;	99US-0148171P.	DB:	3	Gaps:	0
PR 11-AUG-1999;	99US-0148319P.	US-10-067-832D-1 (1-370) x AAG27281 (1-73)			
PR 12-AUG-1999;	99US-0148341P.				
PR 13-AUG-1999;	99US-0148565P.	QY 29 ATGATTCAGGTGGTTTGCACGACCGTCTAGGAAGAAGTCCGGTTAAGTCGACACC 88			
PR 13-AUG-1999;	99US-0148684P.				
PR 16-AUG-1999;	99US-0149336P.	Db 1 MetIleGluValValLeuAsnAspArgLeuGlyLysLysValhrgValysCysaenGlu 20			
PR 17-AUG-1999;	99US-0149175P.				
PR 18-AUG-1999;	99US-0149426P.	QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGCAGCTCGTTGGAAT 148			
PR 20-AUG-1999;	99US-0149722P.				
PR 20-AUG-1999;	99US-0149723P.	Db 21 AspAspThrIleGlyAspLeuAsnLysLeuValalaGlnThrGlyThrArgProGlu 40			
PR 20-AUG-1999;	99US-0149929P.				
PR 23-AUG-1999;	99US-0149930P.	QY 149 AAGATCGTTCCTAAAAGTGTGTACACGATTTTAAAGACCATGATCTCTGGAGATTAT 208			
PR 25-AUG-1999;	99US-0150566P.				
PR 26-AUG-1999;	99US-0150884P.	Db 41 LysIleArgIleGlnLysTyrAsnIleTyrLysAspTyrIleThrLeuLysAspTyr 60			
PR 27-AUG-1999;	99US-0151065P.				
PR 27-AUG-1999;	99US-0151066P.	QY 209 GMAATCCAGATGGATGACCTGGAGCTTTATTAC 244			
PR 27-AUG-1999;	99US-0151080P.				
PR 30-AUG-1999;	99US-0151303P.	Db 61 GluValHisAspGlyMetGlyLeuGluLeuTyrTyr 72			
PR 31-AUG-1999;	99US-0151438P.				
PR 01-SEP-1999;	99US-0151930P.	RESULT 11			
PR 07-SEP-1999;	99US-0152363P.	ADTS8384			
PR 10-SEP-1999;	99US-0153070P.	ID ADTS8384 standard; protein; 73 AA.			
PR 13-SEP-1999;	99US-0153758P.	XX			
PR 15-SEP-1999;	99US-0154018P.	AC ADTS8384;			
PR 16-SEP-1999;	99US-0154039P.	XX			
PR 20-SEP-1999;	99US-0154779P.	DT 13-JAN-2005 (first entry)			
PR 22-SEP-1999;	99US-0155139P.	XX			
PR 23-SEP-1999;	99US-0155486P.	DE Plant polypeptide, SEQ ID 8461.			
PR 24-SEP-1999;	99US-0155659P.	XX			
PR 28-SEP-1999;	99US-0156458P.	KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;			
PR 28-SEP-1999;	99US-01565396P.	KW disease resistance; galactomannan production; plant growth regulator;			
PR 04-OCT-1999;	99US-0157117P.	KW heat tolerance; herbicide tolerance; lignin production;			
PR 05-OCT-1999;	99US-0157753P.	KW extreme osmotic condition tolerance; pathogens resistance;			
PR 06-OCT-1999;	99US-0157865P.	KW pest resistance; yield improvement; seed oil yield; seed protein yield.			
PR 07-OCT-1999;	99US-0158029P.	XX			
PR 08-OCT-1999;	99US-0158232P.	OS Viridiplantae.			
PR 12-OCT-1999;	99US-0158369P.	XX			
PR 13-OCT-1999;	99US-0159293P.	PN US2004216190-A1.			
PR 13-OCT-1999;	99US-0159294P.	XX			
PR 13-OCT-1999;	99US-0159295P.	PD 28-OCT-2004.			
PR 14-OCT-1999;	99US-0159329P.	XX			
PR 14-OCT-1999;	99US-0159330P.	XX			
PR 14-OCT-1999;	99US-0159331P.	PF 18-DEC-2003; 2003US-00739930.			
PR 14-OCT-1999;	99US-0159637P.	XX			
PR 14-OCT-1999;	99US-0159638P.	PR 28-APR-2003; 2003US-00424599.			
PR 18-OCT-1999;	99US-0159594P.	XX			
PR 21-OCT-1999;	99US-0160741P.	XX			
PR 21-OCT-1999;	99US-0160767P.	PA (KOVA/) KOVALIC D K.			
PR 21-OCT-1999;	99US-0160768P.	XX			
PR 21-OCT-1999;	99US-0160770P.	PI Kovalic DK;			
PR 21-OCT-1999;	99US-0160814P.	XX			
PR 21-OCT-1999;	99US-0160815P.	XX			
PR 22-OCT-1999;	99US-0160980P.	XX			
PR 22-OCT-1999;	99US-0160981P.	PT New recombinant DNA constructs useful in the field of biochemistry and			
PR 22-OCT-1999;	99US-0160989P.	PT Genetics, and in particular for producing transgenic plants with improved			
PR 25-OCT-1999;	99US-0161404P.	XX biological characteristics.			
PR 25-OCT-1999;	99US-0161405P.	XX			
PR 25-OCT-1999;	99US-0161406P.	PS Claim 2; SEQ ID NO 8461; 14pp; English.			
PR 26-OCT-1999;	99US-0161359P.	XX			
PR 26-OCT-1999;	99US-0161360P.	CC The invention relates a recombinant DNA construct comprising a			
PR 26-OCT-1999;	99US-0161361P.	CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:			
PR 28-OCT-1999;	99US-0161920P.	CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences			
PR 28-OCT-1999;	99US-0161922P.	CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,			
PR 28-OCT-1999;	99US-0161993P.	CC Arabidopsis, wheat and rape but the specification does not indicate which			
PR 28-OCT-1999;	99US-0162142P.	CC sequences is derived from which organism. Also included is a method of			
		CC producing a plant having an improved property, comprising transforming a			
		CC plant with a recombinant DNA construct comprising a promoter region			
		CC functional in a plant cell operably joined to a polynucleotide encoding a			
		CC polypeptide associated with the property, and growing the transformed			
		CC plant. The property is selected from improving plant cold tolerance, for			
		CC manipulating growth rate in plant cells by modification of the cell cycle			
		CC pathway, for improving plant drought tolerance, for providing increased			
		CC resistance to plant disease, for galactomannan production, for production			

## Alignment Scores:

Pred. No.:	2.4e-27	Length:	73
Score:	300.00	Matches:	55
Percent Similarity:	87.5%	Conservative:	8
Best Local Similarity:	76.4%	Mismatches:	9
Query Match:	45.1%	Indels:	0

CC of plant growth regulators, for improving plant heat tolerance, for  
CC improving plant tolerance to herbicides, for increasing the rate of  
CC homologous recombination in plants, for lignin production, for improving  
CC plant tolerance to extreme osmotic conditions, for improving plant  
CC tolerance to pathogens or pests, for yield improvement by modification of  
CC photosynthesis, for modifying seed oil yield and/or content, for  
CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20040216190.

XX SQ Sequence 73 AA;

Alignment Scores:  
Pred. No.: 1.62e-25 Length: 73  
Score: 285.00 Matches: 56  
Percent Similarity: 86.3% Conservative: 7  
Best Local Similarity: 76.7% Mismatches: 8  
Query Match: 42.9% Indels: 2  
DB: 8 Gaps: 2

US-10-067-832D-1 (1-370) x ADTS8384 (1-73)

QY 29 ATGATTGAGTGGTTTCACACGCGCTAGGAAGAAAGTCGCGTTAAGTGCACACC 88  
DB 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgValLysCys----\*\*\* 19  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGC---ACTCGTTGG 145  
DB 20 AspAspThrileGlyAspLeuLysLeuValAlaAlaGlnThrGly\*\*\*ThrArgPro 39  
QY 146 AATAAGATCGTCTTAAAAAGTGTACACCATTTTAAAGACCATGTATCTCGGAGAT 205  
DB 40 GluLysileArgileGlnLysTrpTyraNileTyryLysAspHisileThrLeuLysAsp 59  
QY 206 TATGAAATCCACCATGGATGGTGAACCTCGGAGCTTTATTAC 244  
DB 60 TyrGluValHisAspGlyMetGlyLeuGluLeuTyrr 72

RESULT 12

ADH32312  
ID ADH32312 standard; protein; 89 AA.

XX AC ADH32312;

XX DT 11-MAR-2004 (first entry)

XX DE Novel yeast smORF572-encoded polypeptide, SEQ ID NO:770.

XX KW Open reading frame; ORF; identification; in silico; yeast; smORF;  
XX NW small open reading frame; antisense therapy; antibody therapy;  
XX KW drug screening; fungal infection; fungicide; gene therapy; vaccine.

XX OS Saccharomyces cerevisiae.

XX PN WO200268693-A2.

XX PD 06-SEP-2002.

XX PF 27-FEB-2002; 2002WO-US005677.

XX PR 27-FEB-2001; 2001US-0271406P.  
XX PR 29-NOV-2001; 2001US-0333726P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Zeng Q, Kessler M, Cottarel G;  
XX DR WPI; 2002-707018/76.  
XX DR N-PSDB; ADH31639.

XX Identifying open reading frames (ORFs) in a genome of an organism by  
XX comparing genomic sequences to one or more genomic libraries containing  
XX ORFs, useful for diagnosing, preventing and/or treating fungal  
XX infections.

XX Claim 38; SEQ ID NO 770; 160pp; English.

XX The invention relates to a method of identifying open reading frames  
XX (ORFs) in a genome of an organism. The method comprises: collecting a  
XX genomic sequence of a first organism; comparing the genomic sequence of  
XX the first organism to one or more other genomic libraries comprising  
XX genomes of other organisms containing ORFs; and determining ORFs for the  
XX first organism based on the comparison. The invention also relates to  
XX smORFs (small open reading frames) from the Saccharomyces cerevisiae  
XX genome which encode polypeptides less than 100 amino acids long; vectors  
XX and host cells comprising a smORF; polypeptides encoded by the smORFs;  
XX antisense compounds targeted to the smORFs and methods for their use in  
XX inhibiting smORF expression; a method of identifying an inhibitor of a  
XX smORF-encoded protein; an antibody against a smORF-encoded polypeptide;  
XX and a composition comprising a smORF or smORF-encoded polypeptide and a  
XX carrier. The invention provides a useful in silico method of identifying  
XX new coding sequences, including homologues of coding sequences, in yeast  
XX and other organisms. The compositions of the invention, and methods of  
XX their use or discovery are also useful for diagnosing, preventing and/or  
XX treating fungal infections. Sequences ADH32216-ADH32888 represent  
XX specifically claimed yeast smORF-encoded polypeptides. Note: The sequence  
XX data for this patent is not represented in the printed specification but  
XX is based on sequence information supplied to Derwent by the European  
XX Patent Office.

XX SQ Sequence 89 AA;

Alignment Scores:  
Pred. No.: 1.8e-21 Length: 89  
Score: 252.00 Matches: 47  
Percent Similarity: 81.9% Conservative: 12  
Best Local Similarity: 65.3% Mismatches: 13  
Query Match: 37.9% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-1 (1-370) x ADH32312 (1-89)

QY 29 ATGATTGAGTGGTTTCACACGCGCTAGGAAGAAAGTCGCGTTAAGTGCACACC 88  
DB 17 MetileGluValValValAenAspArgLeuGlyLysValArgValLysCysLeuAla 36  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGCCTCGTTGAAT 148  
DB 37 GluAspSerValGlyAspPhelysValLysValLeuSerLeuGlnileGlyThrGlnProAen 56  
QY 149 AAGATCGTCTTAAAAAGTGTACACGATTTTAAAGACCATGTATCTCTGGGAGATTAT 208  
DB 57 LysileValLeuGlnLysGlyLysValLeuLysAspHisileSerLeuGluAspTyrr 76  
QY 209 GAAATCCACCATGGATGAACCTCGGAGCTTTATTAC 244  
DB 77 GluValHisAspGlnThrAenLeuGluLeuTyrr 88

RESULT 13

ABR83620  
ID ABR83620 standard; protein; 320 AA.

XX

AC ABR83620;  
 XX 16-OCT-2003 (first entry)  
 DT HUB1-GFP protein sequence SEQ ID NO:11.  
 XX  
 XX Protein expression; purification; fusion protein; SUMO; GFP; RUB; HUB;  
 KW green fluorescent protein; enhancing protein expression level; APG8;  
 KW APG12; URM1; ISG15; secretion.  
 XX  
 XX Homo sapiens.  
 OS Aequorea victoria.  
 OS Synthetic.  
 XX  
 XX WO2003057174-A2.  
 XX 17-JUL-2003.  
 XX  
 XX 07-JAN-2003; 2003WO-US000436.  
 XX  
 XX 07-JAN-2002; 2002US-0346449P.  
 XX (LIFE-) LIFESSENSORS INC.  
 XX  
 XX Butt TR, Weeks SD, Tran HT, Malakhov MP, Malakhova OA;  
 XX WPI; 2003-577500/54.  
 XX N-PSDB; ACF57380.  
 XX  
 XX Enhancing protein expression levels in or from a cell comprises operably  
 PT linking ubiquitin or ubiquitin-like proteins to a nucleic acid encoding a  
 FT protein to generate a construct encoding a fusion protein for expression  
 PT in a host cell.  
 XX  
 XX Example 1; Fig 28A-B; 144pp; English.  
 XX  
 XX The present invention describes a method for enhancing expression levels  
 CC of a protein in or from a cell. The method comprises: (a) operably  
 CC linking a nucleic acid sequence encoding SUMO, RUB, HUB, APG8, APG12,  
 CC URM1 or ISG15, to a nucleic acid sequence encoding the protein to a  
 CC generate a construct encoding a fusion protein; and (b) introducing the  
 CC nucleic acid into a host cell, where the presence of the molecule in the  
 CC fusion protein increases the expression level or the secretion of the  
 CC protein in the host cell. The method is useful for enhancing protein  
 CC expression, and for generating novel amino termini on proteins of  
 CC interest for a variety of research, diagnostic and therapeutic  
 CC applications. The method of enhancing protein circumvents problems  
 CC encountered with previous methods by enhancing the expression of under-  
 CC expressed proteins, increasing the solubility of proteins that are  
 CC insoluble, protecting the candidate proteins from degradation by  
 CC intracellular proteases by fusing ubiquitin-like proteins (UBLs) to their  
 CC N-termini, cleaving the fusion protein to efficiently generate authentic  
 CC proteins using naturally present enzymes, generating proteins with novel  
 CC amino termini, and cleaving all fusion proteins with remarkable  
 CC efficiency irrespective of the N-terminal sequence of the fused protein,  
 CC using UBL hydrolases. The present sequence represents a HUB1-GFP protein  
 CC sequence, which is used in the exemplification of the present invention  
 XX  
 XX Sequence 320 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,53e-21 Length: 320  
 Score: 252.00 Matches: 47  
 Percent Similarity: 81.9% Conservative: 12  
 Best Local Similarity: 65.3% Mismatches: 13  
 Query Match: 37.9% Indels: 0  
 DB: 6 Gaps: 0

US-10-067-832D-1 (1-370) x ABR83620 (1-320)

QY 29 ATGATTCAGGTGGTTTCACGACCGCTAGGAAGAAAGTCGCGTTAGTGCACACACC 88  
 Db 10 MetileGluValValValAsnAspArgLeuGlyLysLysValArgValLysCysLeuAla 29

QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAACTGGCACTCGTTGGAAT 148  
 Db 30 GluAspSerValGlyAspPheLysLysValLeuSerLeuGlnIleGlyThrGlnProhan 49  
 QY 149 AAGATCGTCTTAAAAAGTGTACACGATTTTAAAGACCATGTATCTCTGGAGATTAT 208  
 Db 50 LysIleValLeuGlnLysGlySerValLeuLysAspHieIleSerLeuGluAspTyr 69  
 QY 209 GAATCCACGATGGATGACCTGACCTGAGCTTTATTATAC 244  
 Db 70 GluValHisAspGlnThrAsnLeuGluLeuTyrTyr 81

RESULT 14  
 AAY08414  
 ID AAY08414 standard; protein; 33 AA.  
 XX  
 AC AAY08414;  
 DT 24-JUL-1999 (first entry)  
 XX  
 DE Human beacon protein.  
 XX  
 KW Beacon; hypothalamus; obese; lean; agonist; antagonist; treatment;  
 KW obesity; anorexia; weight maintenance; energy imbalance; diabetes;  
 KW metabolic syndrome; dyslipidemia; hypertension; insulin resistance;  
 KW medicament; livestock; diagnosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 15  
 FT /label= unknown  
 FT /note= "encoded by CNC"  
 XX  
 XX WO99232217-A1.  
 XX  
 XX 14-MAY-1999.  
 XX  
 XX 30-OCT-1998; 98WO-AU000902.  
 XX  
 XX 31-OCT-1997; 97AU-00000117.  
 XX 11-NOV-1997; 97AU-00000323.  
 XX  
 XX (ITDI-) INT DIABETES INST.  
 XX (UYDE-) UNIV DEAKIN.  
 XX  
 XX Zimmet PZ, Collier G;  
 XX WPI; 1999-337484/28.  
 XX N-PSDB; AAX57370.  
 XX  
 XX New gene encoding a beacon protein associated with modulation of obesity,  
 XX diabetes and metabolic energy levels.  
 XX  
 XX Claim 2; Page 54; 85pp; English.  
 XX  
 XX This invention describes a novel beacon protein and its encoding nucleic  
 XX acid which is expressed in larger amounts in hypothalamus tissue of obese  
 XX animals compared to lean animals. Agonists and antagonists of beacon can  
 XX be used to treat obesity, anorexia, weight maintenance, energy imbalance,  
 XX diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin  
 XX resistance. The beacon protein, itself is used to manufacture medicaments  
 XX for treatment of obesity, anorexia, energy imbalance or diabetes. The  
 XX treatment is contemplated for both human and animals, such as those  
 XX important to the livestock industry. The antibody and polynucleotides are  
 XX useful in diagnosis of conditions as above  
 XX  
 XX Sequence 33 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 9.88e-11 Length: 33  
 Score: 163.00 Matches: 32



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:39:29 ; Search time 5.87924 Seconds  
(without alignments)  
1211.048 Million cell updates/sec

Title: US-10-067-832D-1  
Perfect score: 665  
Sequence: 1 gttccaggagattacagctc.....tttgatgctgcacaaaaaaa 370

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat\_07042006\_175613\_9739/app\_query.fasta\_1  
-DB=pir -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=spect -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USER=US10067832 @CGN\_1\_1\_77 -runat\_07042006\_175613\_9739 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 80:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	331	49.8	73	2 T25763	hypothetical prote
2	306	46.0	73	2 T47458	hypothetical prote
3	301	45.3	73	2 T40200	ubiquitin-like pro
4	252	37.9	73	2 S78735	protein YNR032c-a
5	80	12.0	631	2 S55243	upiquitin-like pro
C 6	77	11.6	2606	2 T03159	large tegument pro
C 7	76.5	11.5	274	2 T20786	hypothetical prote
8	74.5	11.2	307	2 T45998	hypothetical prote
9	74.5	11.2	1066	2 B95037	hyaluronidase (imp
10	73.5	11.1	1078	2 F97907	hyaluronate lyase
11	72.5	10.9	299	2 S60529	envelope polyprote
12	72.5	10.9	299	2 S60528	envelope polyprote
C 13	70.5	10.6	662	2 H84984	bo-type ubiquinol
14	70	10.5	551	2 C84549	probable ubiquitin

ALIGNMENTS

RESULT 1

T25763  
hypothetical protein F46F11.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T25763  
R/Pauley, A.; Gattung, S.  
Submitted to the EMBL Data Library, February 1997  
A/Description: The sequence of C. elegans cosmid F46F11.  
A/Reference number: Z20083  
A/Accession: T25763  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-73 <PAU>  
A/Cross-references: UNIPROT:P91302; UNIPARC:UPI0000076B0C; EMBL:U08173; PIDN:AAB42266.1;  
A/Experimental source: strain Bristol N2; clone F46F11  
C/Genetics:  
A/Gene: CESP:F46F11.4  
A/Map position: 1  
A/Introns: 38/2

Alignment Scores:

Pred. No.: 8.72e-29 Length: 73  
Score: 331.00 Matches: 59  
Percent Similarity: 89.0% Conservatives: 6  
Best Local Similarity: 80.8% Mismatches: 8  
Query Match: 49.8% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x T25763 (1-73)

Qy 29 ATGATTGAGTGGTGTTCACACGCGCTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88

Db 1 MetilegluilethrValAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20

Qy 89 GATGACACCATCGGGGACTTGAAGAACTGATACGGCCCAACTGCGACTCGCTTGAAT 148

Db 21 SerAspThrIleGlyAspLeuLysLeuIleAlaAlaGlnThrGlyThrArgtrpGlu 40

```
QY 149 AAGATCGTCTTAAAGTGGTACACGATTTTAAAGACCATGATCTCTGGGAGATTAT 208
  |||
Db 41 LysileValleuLysLysTyrThrLysAspHisIleThrLeuMetAspTyr 60
  |||
QY 209 GAAATCCACCATGGGATGAACCTGGAGCTTTATTACAG 247
  |||
Db 61 GluIleHisGluGlyPheAsnPheGluLeuTyrTyrGln 73

RESULT 2
T47458
hypothetical protein T14D3.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47458
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224467
A:Accession: T47458
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-73 <JOK>
A:Cross-references: UNIPROT:Q9MU1; UNIPARC:UPI000000A81D7; EMBL:AL138649
A:Experimental source: cultivar Columbia; BAC clone T14D3
C:Genetics:
A:Map position: 3
A>Note: T14D3.120

Alignment Scores:
Pred. No.: 5.59e-26 Length: 73
Score: 306.00 Matches: 57
Percent Similarity: 87.5% Conservative: 6
Best Local Similarity: 79.2% Mismatches: 9
Query Match: 46.0% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x T47458 (1-73)

QY 29 ATGATTGAGTGGTTCACAGACCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
  |||
Db 1 MetileGluValLeuAsnAspArgLeuGlyLysLysValArgValLysCysAsnGlu 20
  |||
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
  |||
Db 21 GluAspThrIleGlyAspLeuLysLysValAlaGlnThrGlyThrArgProGlu 40
  |||
QY 149 AAGATCGTCTTAAAGTGGTACACGATTTTAAAGACCATGATCTCTGGGAGATTAT 208
  |||
Db 41 LysileArgileGlnLysTyrAsnIleTyrLysAspHisIleProLeuLysAspTyr 60
  |||
QY 209 GAAATCCACCATGGGATGAACCTGGAGCTTTATTAC 244
  |||
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

RESULT 3
T40200
ubiquitin-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40200
R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, March 1999
A:Reference number: 221912
A:Accession: T40200
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-73 <OLI>
A:Cross-references: UNIPROT:O94650; UNIPARC:UPI000000A678; EMBL:AL049190; PIDN:CAB39137.
A:Experimental source: strain 972h-; cosmid c31E1
C:Genetics:
A:Gene: SPBC31E1.03
A:Map position: 2
A:Introns: 6/2; 40/2
```

```
Alignment Scores:
Pred. No.: 2.04e-25 Length: 73
Score: 301.00 Matches: 53
Percent Similarity: 90.3% Conservative: 12
Best Local Similarity: 73.6% Mismatches: 7
Query Match: 45.3% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x T40200 (1-73)

QY 29 ATGATTGAGTGGTTCACAGACCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
  |||
Db 1 MetileGluValLeuCysAsnAspArgLeuGlyLysLysValArgValLysCysMetPro 20
  |||
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
  |||
Db 21 AspAspThrValGlyAspPheLysLysLeuValAlaGlnThrGlyThrArgProArg 40
  |||
QY 149 AAGATCGTCTTAAAGTGGTACACGATTTTAAAGACCATGATCTCTGGGAGATTAT 208
  |||
Db 41 ArgileValleuLysLysIlePheSerValPheLysAspAsnIleThrLeuAlaAspTyr 60
  |||
QY 209 GAAATCCACCATGGGATGAACCTGGAGCTTTATTAC 244
  |||
Db 61 GluIleHisAspGlyMetSerLeuGluMetTyrTyr 72

RESULT 4
S78735
protein YNR032c-a - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 15-Jan-1999
C:Accession: S78735
R:Fohl, I.M.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63346
A:Accession: S78735
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-73 <POH>
A:Cross-references: UNIPARC:UPI0000168403; EMBL:Z71647; MIPS:YNR032c-a
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 14R

Alignment Scores:
Pred. No.: 6.45e-20 Length: 73
Score: 252.00 Matches: 47
Percent Similarity: 81.9% Conservative: 12
Best Local Similarity: 65.3% Mismatches: 13
Query Match: 37.9% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x S78735 (1-73)

QY 29 ATGATTGAGTGGTTCACAGACCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
  |||
Db 1 MetileGluValValLeuAsnAspArgLeuGlyLysLysValArgValLysCysLeuAla 20
  |||
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
  |||
Db 21 GluAspSerValGlyAspPheLysLysValLeuSerLeuGlnIleGlyThrGlnProAsn 40
  |||
QY 149 AAGATCGTCTTAAAGTGGTACACGATTTTAAAGACCATGATCTCTGGGAGATTAT 208
  |||
Db 41 LysileValleuGlnLysGlySerValLeuLysAspHisIleSerLeuGluAspTyr 60
  |||
QY 209 GAAATCCACCATGGGATGAACCTGGAGCTTTATTAC 244
  |||
Db 61 GluValHisAspGlnThrAsnLeuGluLeuTyrTyr 72

RESULT 5
S55243
```

ubiquitin-like protein 8 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C>Date: 10-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Oct-2004  
C:Accession: S55243; S61068  
R:Callis, J.; Carpenter, T.; Sun, C.W.; Vierstra, R.D.  
Genetics 139, 921-939, 1995  
A:Title: Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins  
A:Reference number: S55243; PMID:95229071; PMID:7713442  
A:Accession: S55243  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-631 <CAL>  
A:Cross-references: UNIPROT:Q39256; UNIPARC:UPI000017A48A; EMBL:L05917  
A:Experimental source: ecotype Columbia  
R:Callis, J.; Carpenter, T.; Sun, C.W.; Vierstra, R.D.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S61068  
A:Accession: S61068  
A:Molecule type: DNA  
A:Residues: 1-341, 'E', 343-631 <CAW>  
A:Cross-references: UNIPARC:UPI0000A3588; EMBL:L05917; NID:g870793; PIDN:AAA68879.1; PI  
F:73-78/Domain: ubiquitin homology <UBH1>  
F:79-154/Domain: ubiquitin homology <UBH2>  
F:155-237/Domain: ubiquitin homology <UBH3>  
F:238-318/Domain: ubiquitin homology <UBH4>  
F:319-392/Domain: ubiquitin homology <UBH5>  
F:393-468/Domain: ubiquitin homology <UBH6>  
F:469-551/Domain: ubiquitin homology <UBH7>  
F:552-627/Domain: ubiquitin homology <UBH8>

## Alignment Scores:

Pred. No.:	1.39	Length:	631
Score:	80.00	Matches:	19
Percent Similarity:	58.3%	Conservative:	16
Best Local Similarity:	31.7%	Mismatches:	25
Query Match:	12.0%	Indels:	0
DB:	2	Gaps:	0

US-10-067-832D-1 (1-370) x S55243 (1-631)

Qy	59	GGAAAGAAAGTCGGGTTAAGTGCACACGATGACACACGCGGACTTGAAGAACTG	118
Db	561	GlyysThrIleLeuGluValGluSerSerThrIleAlaAsnValysGluLys	580
Qy	119	ATAGCGGCCCAACTCGCTGGTGAATAGATGCTTCTTAAAGAGTGGTACACGATT	178
Db	581	IleGlnValysGluGlyIleLeuProAspGlnMetLeuIlePhePheGlyGlnGln	600
Qy	179	TTTAAGGACCATGATCTCTGGGAGATTATGAATCCAGATGGATGAACCTGGAGCTT	238
Db	601	LeuGluAspGlyValThrLeuGlyAspThrAspIleHisLysLysSerThrLeuTyrl	620

## RESULT 6

T03159  
large tegument protein - alcelaphine herpesvirus 1  
C:Species: alcelaphine herpesvirus 1  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T03159  
R:Esser, A.; Pflanz, R.; Fleckenstein, B.  
J. Virol. 71, 6517-6525, 1997  
A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.  
A:Reference number: Z14840; MUID:97404659; PMID:9261371  
A:Accession: T03159  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2606 <ENS>  
A:Cross-references: UNIPROT:Q36414; UNIPARC:UPI00000F494A; EMBL:AF005370; NID:g2337967;

## Alignment Scores:

Pred. No.:	3.12	Length:	2606
Score:	77.00	Matches:	32
Percent Similarity:	42.6%	Conservative:	11
Best Local Similarity:	31.7%	Mismatches:	30

## Query Match:

DB:	11.6%	Indels:	28
	2	Gaps:	6

US-10-067-832D-1 (1-370) x T03159 (1-2606)

Qy	256	TCCCTCTACTGGTAAATAAGCTCCAGGTCATCCCATCGTGATTTTCATATCTCCAG	197
Db	1724	AlaProThrGlnProIleThrLeuGlnValPheProThrValAsp---AsnLeuSerMet	1742
Qy	196	AGATACATGGTCTTAAATAATCGTGACCACTT-----TTTAAGAACGATCTT	149
Db	1743	AspThrIleProIleLysAsnAlaSerProLeuSerLeuGlnValIlePheAsnAsnPhe	1762
Qy	148	ATTCAACAGAGTCCAGGTTTGGCGGTATCAGTTTCTTCAAGTCCCGATGGT-----	95
Db	1763	IleGluThr-----TyrPheValGlnAlaProGlnGlyProGln	1775
Qy	94	-----GTCATCGGTGTTCACCTTAAACGCGGACTTT	65
Db	1776	LysAspThrSerGlnTyrArgGlySerThrValLeuProAsnValLeuGlnAlaHisPhe	1795
Qy	64	CTTTCCTAGAGCGGTGTTGCAACCCACCTCAATCATTTGTTGGTGGAGCTGTAATCTCCTG	5
Db	1796	-----GlyThrIleAlaThrHisLeuIleAsnSerHisTrpAsn---AsnIleLeu	1811
Qy	4	GAA	2
Db	1812	Glu	1812

## RESULT 7

T20786

hypothetical protein F11B6.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T20786  
R:Matthews, L.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19324  
A:Accession: T20786  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-274 <WIL>  
A:Cross-references: UNIPROT:Q9XVQ9; UNIPARC:UPI0000033PF6; EMBL:Z81058; PIDN:CAB02921.1;  
A:Experimental source: clone F11B6  
C:Genetics:  
A:Gene: CESP:F11B6.5  
A:Map position: 4  
A:Introns: 38/3; 99/1; 183/2; 235/2

## Alignment Scores:

Pred. No.:	3.36	Length:	274
Score:	76.50	Matches:	25
Percent Similarity:	41.4%	Conservative:	21
Best Local Similarity:	22.5%	Mismatches:	42
Query Match:	11.5%	Indels:	23
DB:	2	Gaps:	3

US-10-067-832D-1 (1-370) x T20786 (1-274)

Qy	370	TTTTTTTTCGACGATCAAGTTTATTTCATGTAATGTTTAAATAGCATCTACAA	311
Db	146	TrpTyrHisHisAlaLeuThrPheValThrValAlaValThrTyrSerGluHis-----	163
Qy	310	CAGTGTTAAATGAGCCATCGGAGAGAAAGAGGTTGGGCAAGGTGGAGGAATTCCTCCT	251
Db	164	-----GlnAlaTrpAlaArgTrpSer-----	170
Qy	250	CTACTGTGTAATAAGTCCAGGTTTCATCCCATCGTGGATTTTCATAATCTCCAGAGATAC	191
Db	171	---LeuAlaLeuAsnLeuAlaValHisThrValMetTyrPheTyrPheAlaValArgala	189
Qy	190	ATGTCCTCTTAAATCGTGATCCACTTTTAAAG-----AACGAT	152

Db 190 LeuAsnIleGlnThrProArgProValAlaLysPheIleThrThrIleGlnIleValGln 209  
QY 151 CTTATTCCACAGAGTGCAGTTGGGCGGTATCAGTTTCTTCAAGTCCCGATGGTGTC 92  
Db 210 PheValIleSerCysTyrIlePheGlyHisLeuValPheIleLysSerAlaAspSerVal 229  
QY 91 ATCGGTGTCACCTTAACGGCGGACTTCTTTTCC 59  
Db 230 ProGlyCysAlaValSerTrpAsnValLeuSer 240  
RESULT 8  
T45998  
hypothetical protein F9D24.290 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T45998  
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: 223011  
A;Accession: T45998  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-307 <DAN>  
A;Cross-references: UNIPROT:Q9M2H8; UNIPARC:UPI000009B50F; EMBL:AL137081  
A;Experimental source: cultivar Columbia; BAC clone F9D24  
C;Genetics:  
A;Map position: 3  
A;Introns: 36/2; 90/1; 170/3  
A;Note: F9D24.290  
C;Superfamily: Arabidopsis thaliana hypothetical protein F9D24.110  
Alignment Scores:  
Pred. No.: 5.65 Length: 307  
Score: 74.50 Matches: 22  
Percent Similarity: 40.8% Conservative: 9  
Best Local Similarity: 28.9% Mismatches: 24  
Query Match: 11.2% Indels: 21  
DB: 2 Gaps: 2  
US-10-067-832D-1 (1-370) x T45998 (1-307)  
QY 116 CTGATAGCGGCCAACTGGCACTCGTTGGTAAGATCGTTCTTAAAGTGGTACACG 175  
Db 30 LeuIleAla-----GlySerArgTyrArgLeuLeuAlaLeuProLys----- 43  
QY 176 ATTTTAAAGACCATGTATCTCTGGGAGATTATGAATCCACGATGGGATGAACCTGGAG 235  
Db 44 -----GlyThrAsnTyrGlu 48  
QY 236 CTTTATTACCAGTAGAGGGAATTCCTCCACCTTGCCCAACCTTGCTTCTCTCCCATG 295  
Db 49 PhePheTyrGlnTyrMetGlyValAlaAspSerCysGlnSerLeuThrSerSerTrpArg 68  
QY 296 GCTCATTAAACACTGTGTAGATGCTCATTTTAAACAATTCACATGAA 343  
Db 69 ArgHisValLysLeuArgLeuThrIleValAsnGlyIleSerHisLys 84  
RESULT 9  
B95037  
hyaluronidase [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: B95037  
R;Retzlaff, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
n, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: B95037  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-1066 <KUR>  
A;Cross-references: UNIPROT:Q54873; UNIPARC:UPI000012D052; GB:AE005672; PIDN:AAK74491.1;  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0314  
Alignment Scores:  
Pred. No.: 5.83 Length: 1066  
Score: 74.50 Matches: 23  
Percent Similarity: 46.3% Conservative: 8  
Best Local Similarity: 34.3% Mismatches: 27  
Query Match: 11.2% Indels: 9  
DB: 2 Gaps: 3  
US-10-067-832D-1 (1-370) x B95037 (1-1066)  
QY 65 AAAGTCGGCTTAAGTGCACACCGATGACACCATCGGGACTTGAAGAACTGATAGCG 124  
Db 128 LysLeuArgPheLysIleLysThrAspAsnLysIleGlyIleAlaLysValArgIleIle 147  
QY 125 GCCCAAACTGGC-----ACTCGTTGGATAAGATCGTT-----CTTAAAAAG 166  
Db 148 GluGluSerGlyLysAspLysArgLeuTrpAsnSerAlaThrThrSerGlyThrLysAsp 167  
QY 167 TGGTACACGATTTTAAAGACCATGTATCTCTGGGAGATTATCAAAATCCACGATGGGATG 226  
Db 168 TrpGlnThrIleGluAlaAspTyrSerProThrLeuAspVal-----AspLysIle 184  
QY 227 AACCTCGAGCTTTATTACCAG 247  
Db 185 LysLeuGluLeuPheTyrGlu 191  
RESULT 10  
F97907  
hyaluronate lyase (EC 4.2.2.1) [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: F97907  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Es  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: F97907  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1078 <KUR>  
A;Cross-references: UNIPROT:Q8CWU3; UNIPARC:UPI000008E4928; GB:AE007317; PIDN:AAK99090.1;  
C;Genetics:  
A;Gene: hysA  
C;Keywords: carbon-oxygen lyase  
Alignment Scores:  
Pred. No.: 7.55 Length: 1078  
Score: 73.50 Matches: 22  
Percent Similarity: 46.3% Conservative: 9  
Best Local Similarity: 32.8% Mismatches: 27  
Query Match: 11.1% Indels: 9  
DB: 2 Gaps: 3  
US-10-067-832D-1 (1-370) x F97907 (1-1078)  
QY 65 AAAGTCGGCTTAAGTGCACACCGATGACACCATCGGGACTTGAAGAACTGATAGCG 124  
Db 140 LysLeuArgPheLysIleLysThrAspAsnLysValGlyIleAlaLysValArgIleIle 159  
QY 125 GCCCAAACTGGC-----ACTCGTTGGATAAGATCGTT-----CTTAAAAAG 166  
Db 160 GluGluSerGlyLysAspLysArgLeuTrpAsnSerAlaThrThrSerGlyThrLysAsp 179  
QY 167 TGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTATCAAAATCCACGATGGGATG 226



Db 180 TrpGlnThrIleGluAlaAspTyrSerProThrLeuAspVal-----AspLysIle 196  
Qy 227 AACCTGGAGCTTTATTACCAG 247  
Db 197 LysLeuGluLeuPheTyrGlu 203  
RESULT 11  
S60529  
envelope polyprotein gp1 - human immunodeficiency virus type 1 (isolate CI-14-13) (frag  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: isolate CI-14-13  
C;Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S60529  
R;Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Franssen, K.; Motte, J.;  
AIDS 8, 21-26, 1994  
A;Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d  
A;Reference number: S60521; MUID:94280700; PMID:8011235  
A;Accession: S60529  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-299 <JAN>  
C;Cross-references: UNIPROT:Q76172; UNIPARC:UPI000010385B; EMBL:X72031; NID:g468637; PID  
A;Experimental source: isolate CI-14-13  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; glycoprotein; immunodeficiency; polyprotein  
Alignment Scores:  
Pred. No.: 9.48 Length: 299  
Score: 72.50 Matches: 28  
Percent Similarity: 40.4% Conservative: 12  
Best Local Similarity: 28.3% Mismatches: 26  
Query Match: 10.9% Indels: 33  
DB: 2 Gaps: 4  
US-10-067-832D-1 (1-370) x S60529 (1-299)  
Qy 5 CAGGAGATTACAGCTCCAGCCACAAATGATTGAGTGTTTGCACAGCCGCTAGGAAAG 64  
Db 8 GluAsnIleThrAsnAsnAlaLysThrIleVal-----GlnLeuValLys 23  
Qy 65 AAAGTCGCGTTAAGTGAAC----- 85  
Db 24 ProValArgIleAsnCysThrArgProSerAsnAsnThrArgLysSerValProIleGly 43  
Qy 86 -----ACCGATGACACCATCGGGACTTGAGAACTGATAGCG 124  
Db 44 ProGlyGlnAlaPheTyrAlaThrAspAspIleIleGlyAspIleArgGlnAlaHisCys 63  
Qy 125 GCCCAAACTGGCACTCGTTGGAATAAGATCGTT-----CTTAAAG 166  
Db 64 AsnValSerLysThrGluTrpAsnGluThrLeuArgGlnValAlaThrGlnLeuLysLys 83  
Qy 167 TGGTAC-----ACGATTTTAAAGACCATGTATCTCTGGAGATTATGAATC 214  
Db 84 HisPheLysAsnAlaThrIleIlePheAlaAsnProSerGlyGlyAspLeuGluVal 102  
RESULT 12  
S60528  
envelope polyprotein gp1 - human immunodeficiency virus type 1 (isolate CI-14-5 and iso  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: isolate CI-14-5; isolate CI-14-21  
C;Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S60528; S60530  
R;Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Franssen, K.; Motte, J.;  
AIDS 8, 21-26, 1994  
A;Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d  
A;Reference number: S60521; MUID:94280700; PMID:8011235  
A;Accession: S60528  
A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
A;Residues: 1-299 <JAN>  
C;Cross-references: UNIPROT:Q76171; UNIPARC:UPI0000101A6F; EMBL:X72030; PIDN:CAAS0915.1; PI  
A;Experimental source: isolate CI-14-5  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993  
A;Accession: S60530  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-299 <JAN>  
C;Cross-references: UNIPARC:UPI0000101A6F; EMBL:X72032; NID:g468639; PIDN:CAAS0915.1; PI  
A;Experimental source: isolate CI-14-21  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; glycoprotein; immunodeficiency; polyprotein  
Alignment Scores:  
Pred. No.: 9.48 Length: 299  
Score: 72.50 Matches: 28  
Percent Similarity: 40.4% Conservative: 12  
Best Local Similarity: 28.3% Mismatches: 26  
Query Match: 10.9% Indels: 33  
DB: 2 Gaps: 4  
US-10-067-832D-1 (1-370) x S60528 (1-299)  
Qy 5 CAGGAGATTACAGCTCCAGCCACAAATGATTGAGTGTTTGCACAGCCGCTAGGAAAG 64  
Db 8 GluAsnIleThrAsnAsnAlaLysThrIleVal-----GlnLeuValLys 23  
Qy 65 AAAGTCGCGTTAAGTGAAC----- 85  
Db 24 ProValArgIleAsnCysThrArgProSerAsnAsnThrArgLysSerValProIleGly 43  
Qy 86 -----ACCGATGACACCATCGGGACTTGAGAACTGATAGCG 124  
Db 44 ProGlyGlnAlaPheTyrAlaThrAspAspIleIleGlyAspIleArgGlnAlaHisCys 63  
Qy 125 GCCCAAACTGGCACTCGTTGGAATAAGATCGTT-----CTTAAAG 166  
Db 64 AsnValSerLysThrGluTrpAsnGluThrLeuArgGlnValAlaThrGlnLeuLysLys 83  
Qy 167 TGGTAC-----ACGATTTTAAAGACCATGTATCTCTGGAGATTATGAATC 214  
Db 84 HisPheLysAsnAlaThrIleIlePheAlaAsnProSerGlyGlyAspLeuGluVal 102  
RESULT 13  
H84984  
bo-type ubiquinol oxidase (EC 1.10.3.-) chain I [similarity] - Buchnera sp. (strain APS)  
C;Species: Buchnera sp.  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 10-May-2001  
C;Accession: H84984  
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A.  
F;106,421/Binding site: heme a iron (His) (axial ligands) #status predicted  
F;284-288/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted  
F;288/Binding site: oxygen (Tyr) #status predicted  
A;Status: preliminary  
A;Accession: H84984  
A;Molecule type: DNA  
A;Residues: 1-662 <STO>  
C;Cross-references: UNIPARC:UPI000005E5C3; GB:AP0000398; GSPDB:GN00144  
A;Experimental source: strain APS  
C;Genetics:  
A;Gene: cyoB; BU471  
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology  
C;Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex  
F;106,421/Binding site: heme a iron (His) (axial ligands) #status predicted  
F;284-288/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted  
F;288/Binding site: oxygen (Tyr) #status predicted  
Alignment Scores:  
Pred. No.: 16.2 Length: 662



## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:32:23 ; Search time 34.178 Seconds  
(without alignments)  
1527.566 Million cell updates/sec

Title: US-10-067-832D-1

Perfect score: 665

Sequence: 1 gttccaggagattacagctc.....tttgatgctgcacaaaaaaa 370

Scoring table:

XGapop 10.0 , Xgapext 0.5  
 Ygapop 10.0 , Ygapext 0.5  
 Fgapop 6.0 , Fgapext 7.0  
 Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
 -Q=/abs/ABSSWEB/spool/US10067832/runat\_07042006\_175611\_9700/app\_query.fasta\_1  
 -DB=Uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
 -UNITS=bits -SPART=1 -END=-1 -NATRIX=blosum62 -TRANS=human40 cdi -LIST=45  
 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
 -USR=US10067832 @CGN\_1\_580 @runat\_07042006\_175611\_9700 -NCPU=6 -ICPU=3  
 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DRV TIMEOUT=120  
 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	59.2	73	1	UBL5_HUMAN Q9bz11 homo sapien
2	394	59.2	73	1	UBL5_MESAU O6egx7 mesocricetu
3	394	59.2	73	1	UBL5_MOUSE Q9epv8 mus musculus
4	394	59.2	73	1	UBL5_PSAOB Q791b0 psammomye o
5	394	59.2	73	2	QSRC10_FONPY Q9rc10 pongo pygma
6	394	59.2	73	2	Q4R5J3_MACFA Q4r5j3 macaca fasc
7	368	55.3	73	1	UBL5_BRACE Q7sxf2 brachydanio
8	367	55.2	73	2	Q5HZC1_XENLA Q5hzcl xenopus lae
9	354	53.2	73	2	Q4SM16_TETNG Q4sm16 tetraodon n
10	344	51.7	73	2	Q7QHW2_ANOGA Q7qhw2 anopheles g
11	339	51.0	73	1	UBL5_DROME Q9v998 drosophila
12	331	49.8	73	1	UBL5_CAEEL P91302 caenorhabdi
13	331	49.8	73	2	Q617M9_CAEBR Q617m9 caenorhabdi
14	329	49.5	73	2	Q5B857_SCHUA Q5bs57 schistosoma
15	325	48.9	73	2	Q4VAF8_MOUSE Q4vaf8 mus musculus
16	315	47.4	73	2	Q6K220_ORYSA Q6k220 oryza sativ

17	313	47.1	73	1	UBL5_ARATH Q9Fg29 arabidopsis
18	313	47.1	73	2	Q570V8_ARATH Q570v8 arabidopsis
19	308	46.3	73	2	Q7RBM4_PLAYO Q7rbm4 plasmodium
20	308	46.3	73	2	Q4YUJ7_PLABE Q4yuj7 plasmodium
21	306	46.0	73	2	Q9M1U1_ARATH Q9m1u1 arabidopsis
22	302	45.4	73	2	Q8I545_PLAF7 Q8i545 plasmodium
23	301	45.3	73	1	HUB1_SCHPO Q94650 schizosacch
24	293	44.1	73	2	Q4MZ69_THEPA Q4mz69 theileria p
25	291	43.8	76	2	Q4UBB9_CRYNE Q4ubeb9 theileria a
26	288	43.3	276	2	Q55M54_CRYNE Q55m54 cryptococcu
27	288	43.3	276	2	Q5K8L5_CRYNE Q5k8l5 cryptococcu
28	266	40.0	73	1	HUB1_D8HA Q6bup7 debaryomyce
29	263	39.5	70	2	Q6CI04_YARLI Q6ci04 yarrowia li
30	250	37.6	47	2	Q5BJS1_RAT Q65j81 rattus norv
31	248	37.3	73	1	HUB1_YEAST Q6G546 saccharomyc
32	243	36.5	74	1	HUB1_CANGA Q6f1x7 candida gla
33	238.5	35.9	74	1	HUB1_KLJLA Q6cu12 kluyvaromyc
34	238.5	35.9	79	2	Q6CU11_KJULA Q6cu11 kluyvaromyc
35	235	35.3	73	1	HUB1_ASHGO Q756x3 ashbya goss
36	231	34.7	87	2	Q54Q03_DICDI Q54q03 dictyosteli
37	218	32.8	79	2	Q5OP03_ENTHI Q5op03 entamoeba h
38	168	25.3	73	2	Q4P9W2_USTMA Q4p9w2 ustilago ma
39	154.5	23.2	237	2	Q4I016_GIBZE Q4i016 gibberella
40	141.5	21.3	616	2	Q5BCG2_SWENI Q5bcg2 aspergillus
41	129.5	19.5	239	2	Q526Z0_MAGGR Q526z0 magnaporthe
42	119.5	18.0	261	2	Q8X065_NEUCR Q8x065 neurospora
43	116	17.4	189	2	Q4WYK6_ASPFU Q4wyk6 aspergillus
44	98.5	14.8	851	2	Q7ZMI4_9HIV1 Q7zmi4 human immun
45	95.5	14.4	115	2	Q7ZCU2_9HIV1 Q7zcu2 human immun

#### ALIGNMENTS

RESULT 1

IDL	UBL5_HUMAN	STANDARD;	PRT;	73 AA.
AC	Q9BZL1;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Ubiquitin-like protein 5.			
GN	Name=UBL5;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.			
RC	TISSUE=Iris;			
RX	MEDLINE=21100898; PubMed=1161819; DOI=10.1006/geno.2000.6439;			
RA	Friedman J.S., Koop B.F., Raymond V., Walter M.A.;			
RT	"Isolation of a ubiquitin-like (UBL5) gene from a screen identifying			
RT	highly expressed and conserved iris genes.";			
RL	Genomics 71:252-255(2001).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RA	Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,			
RA	Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,			
RA	Phelan M., Farmer A.;			
RT	"Cloning of human full-length cDNAs in BD Creator(TM) system donor			
RT	vector.";			
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Kidney;			
RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426030999;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,			



DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Ubiquitin-like protein 5.  
 GN Name=UBL5;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muroidae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Pancreas, Spinal ganglion, and Tongue;  
 RX MEDLINE=22354683; PubMed=12477932; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusci V., Choctia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Negashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verdardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- SUBUNIT: Interacts with CLK4 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 DR EMBL; AK002730; BAB22312.1; -; mRNA.  
 DR EMBL; AK003992; BAB23111.1; -; mRNA.  
 DR EMBL; AK007726; BAB25215.1; -; mRNA.  
 DR EMBL; AK009854; BAB26545.1; -; mRNA.  
 DR EMBL; AK012803; BAB28481.1; -; mRNA.  
 DR EMBL; AK051149; BAC34537.1; -; mRNA.  
 DR EMBL; BC028498; AAC28498.1; -; mRNA.  
 DR PDB; 1UH6; NMR; A=1-73.  
 DR MGI; MGI:1913427; Ubl5.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PROSITE; PS00299; UBIQUITIN\_1; FALSE NEG.  
 DR PROSITE; PS00553; UBIQUITIN\_2; FALSE NEG.  
 KW 3D-structure; Ubl conjugation pathway.  
 FT DOMAIN 1 73 Ubiquitin-like.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;  
 Alignment Scores:  
 Pred. No.: 1.03e-37 Length: 73  
 Score: 394.00 Matches: 73  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 59.2% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-067-832D-1 (1-370) x UBL5\_MOUSE (1-73)  
 Qy 29 ATGATTCAGTGGTTCACAGCCGCTAGGAAGAAGTCCGCTTAAGTCGACACACC 88  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20  
 Qy 89 GATCACACCATCGGGGACTTGAACAACTGATACGCCCAACTGGCAGCTCGTTGGAAAT 148  
 Db 21 AspAspThrileGlyAspLeuLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
 Qy 149 AAGATCGTTCTTAAAGAGTGGTACAGCATTTTAAAGACCATGATCTCTCGGAGATTAT 208  
 Db 41 LysileValLeuLysLysTrpTyrThrilePheLysAspHisValSerLeuGlyAspTyr 60  
 Qy 209 GAATCCAGTCGATGACCTGGAGCTTATTTATTCACG 247  
 Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyrTrpGln 73  
 RESULT 4  
 UBL5\_PSAOB STANDARD; PRT; 73 AA.  
 ID \_UBL5\_PSAOB  
 AC Q791B0;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Ubiquitin-like protein 5 (Beacon protein).  
 GN Name=UBL5; obesus (Fat sand rat).  
 OS Psammomys obesus (Fat sand rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muroidae; Muridae; Gerbillinae; Psammomys.  
 OX NCBI\_TaxID=48139;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=20527879; PubMed=11078442;  
 RA Collier G.R., McMillan J.S., Windmill K., Walder K., Tenne-Brown J.,  
 RA de Silva A., Trevaaskis J., Jones S., Morton G.J., Lee S., Augert G.,  
 RA Civitarese A., Zimet P.Z.;  
 RT "Beacon: a novel gene involved in the regulation of energy balance.";  
 RL Diabetes 49:1766-1771(2000).  
 CC -1- SUBUNIT: Interacts with CLK4 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AF318186; AAC34704.1; -; mRNA.  
 CC SMR; Q791B0; 1-73.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00503; UBIQUITIN\_2; FALSE\_NEG.  
 KW Ubl conjugation pathway.  
 FT DOMAIN 1 73 Ubiquitin-like.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:  
 Pred. No.: 1.03e-37 Length: 73  
 Score: 394.00 Matches: 73  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 59.2% Indels: 0  
 DB: 1 Gaps: 0

US-10-067-832D-1 (1-370) x UBL5\_PSAOB (1-73)

QY 29 ATGATTGAGTGGTTTGCACACGCGTCTAGGAAGAAAGTCCGCTTAAGTCAACACC 88  
 |||||  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallYsCysAsnThr 20  
 |||||  
 QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148  
 |||||  
 Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAlaGlnThrGlyThrArgTPrasn 40  
 |||||  
 QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGAGATTAT 208  
 |||||  
 Db 41 LysileValLeuLysLysTrpTyrThrIlePheLysAspHisValSerLeuGlyAspTyr 60  
 |||||  
 QY 209 GAAATCCAGCATGGGATGAACCTGGAGCTTTATTACCAG 247  
 |||||  
 Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyrTyrGln 73  
 |||||

RESULT 5

QSRCIO\_PONPY  
 ID QSRCIO\_PONPY PRELIMINARY; PRT; 73 AA.  
 AC QSRCIO;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DE Hypothetical protein DKF2p469G145.  
 GN Names=DKF2p469G145;  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
 OC Pongo.  
 OX NCBI\_TaxID=9600;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RG The German cDNA Consortium;  
 RA Ansoorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR858290; CAH90527.1; -; mRNA.  
 DR SNR; Q5RCIO; 1-73.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:  
 Pred. No.: 1.03e-37 Length: 73

Score: 394.00 Matches: 73  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 59.2% Indels: 0  
 DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x Q5RCIO\_PONPY (1-73)

QY 29 ATGATTGAGTGGTTTGCACACGCGTCTAGGAAGAAAGTCCGCTTAAGTCAACACC 88  
 |||||  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallYsCysAsnThr 20  
 |||||  
 QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148  
 |||||  
 Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAlaGlnThrGlyThrArgTPrasn 40  
 |||||  
 QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGAGATTAT 208  
 |||||  
 Db 41 LysileValLeuLysLysTrpTyrThrIlePheLysAspHisValSerLeuGlyAspTyr 60  
 |||||  
 QY 209 GAAATCCAGCATGGGATGAACCTGGAGCTTTATTACCAG 247  
 |||||  
 Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyrTyrGln 73  
 |||||

RESULT 6

Q4R5J3\_MACFA  
 ID Q4R5J3\_MACFA PRELIMINARY; PRT; 73 AA.  
 AC Q4R5J3;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DE Brain cDNA, clone: OFIA-12296, similar to human ubiquitin-like 5 (UBL5), (Testis cDNA, clone: QrsA-15893, similar to human ubiquitin-like 5 (UBL5)).  
 DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 OC Cercopitheidae; Cercopitheciniae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA International consortium for macaque cDNA sequencing, analysis;  
 RT "DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications."  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,  
 RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;  
 RT "Substitution rate and structural divergence of 5'UTR evolution: Comparative analysis between human and cynomolgus monkey cDNAs."  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB169550; BAE01632.1; -; mRNA.  
 DR EMBL; AB169933; BAE01034.1; -; mRNA.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:  
 Pred. No.: 1.03e-37 Length: 73  
 Score: 394.00 Matches: 73  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 59.2% Indels: 0  
 DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x Q4R5J3\_MACFA (1-73)

QY 29 ATGATTGAGTGGTTTGCACACGCGTCTAGGAAGAAAGTCCGCTTAAGTCAACACC 88  
 |||||  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallYsCysAsnThr 20  
 |||||  
 QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148  
 |||||  
 Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAlaGlnThrGlyThrArgTPrasn 40  
 |||||



29	ATGATTGAGGTGGTTTGCACGACCGCTTAGGAAGAAAGTCCGCGTTTAAGTCACACC	88
1	MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnGln	20
89	GATGACACCATCGGGACTTGAAGAAACTGATAGCGGCCCAACTGGCACTGTTGGAAAT	148
	:::	
21	GluAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpGlu	40
149	AAGATCGTTCCTTAAAAAGTGGTACACCATTTTAAAGGACCATGTACTCTGGAGATTAT	208
	:::	
41	LysIleValLeuLysLysTrpTyrThrIlePheLysAsnHisValSerLeuGlyAspTyr	60
209	GAATTCACGATGGATGAACCTGGAGCTTTATTATC	244
61	GluIleHisAspGlyMetAsnLeuGluLeuTyrTyr	72

US-10-067-832D-1 (1-370) x Q5HZC1\_XENLA (1-73)

	Qy	Dy
29	ATGATTGAGGTGGTTTGCAACACCAGCTCTAGGAAGAAGTCCGCGTTAAGTCGAACACC	88
1	MetIleGluValIValCysAsnAspArgLeuGluValValVaIArgValILvscVssAnSer	20



ID UBL5 DROME STANDARD; PRT; 73 AA.  
 AC Q9V998; Q9V998; 73 AA.  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Ubiquitin-like protein 5.  
 GN Name=l(2)k03203; ORFNames=CG3450;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Fodor C., Gabriellista A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.I., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lanko P., Lei Y., Leick A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusakern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E.E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassaarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RN Science 287:2185-2195 (2000).  
 RN [2]  
 RP GENOME REANNOTATION.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Mirza S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kamikawa J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupay J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review."  
 RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=Berkeley;  
 RA Stapleton M., Carlson J.W., Chavez C., Friese E., George R.A.,  
 RA Pacle J.M., Park S., Wan K.H., Yu C., Celniker S.E.,  
 RN Submitted (JUN-2005) to the ENBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; AE003790; AAF57398.1; -; Genomic\_DNA.  
 CC EMBL; BT023690; AAY85090.1; -; mRNA.  
 CC SMR; Q9V998; 1-73.  
 CC Ensembl; CG3450; Drosophila melanogaster.  
 CC Flybase; FBgn0022224; l(2)k03203.  
 CC GO; GO:0005737; C:cytoplasm; ISS.  
 CC InterPro; IPR000626; Ubiquitin.  
 CC Pfam; PF00240; Ubiquitin; 1.  
 CC PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
 CC PROSITE; PS0053; UBIQUITIN\_2; 1.  
 KW Ub1 conjugation pathway.  
 FT DOMAIN 1 73 Ubiquitin-like.  
 SQ SEQUENCE 73 AA; 8570 MW; 0870111AC586A70 CRC64;  
 Alignment Scores:  
 Pred. No.: 3,79e-31 Length: 73  
 Score: 339.00 Matches: 63  
 Percent Similarity: 90.4% Conservative: 3  
 Best Local Similarity: 86.3% Mismatches: 7  
 Query Match: 51.0% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-067-832D-1 (1-370) x UBL5\_DROME (1-73)  
 QY 29 ATGATTGAGGTGGTTCGACACGCGCTAGAGAAAGAAAGTCGCGTTAAGTGCACACACC 88  
 DB 1 MetileGluLeuThrCysAsnAspArgLeuGlyLysValArgValLysCysAsnPro 20  
 QY 89 GATGACACCATCGGGAGCTTGAAGAACTGATAGCGGCCCAACTGGCAGCTCGTTGGAAAT 148  
 DB 21 AspAspThrileGlyAspLeuLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
 QY 149 AAGATCGTCTTAAAGAGTGTACACCATGTTTAAAGGACCATGATCTCTCGGAGATTAT 208  
 DB 41 LysileValleuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60  
 QY 209 GAAATCCACCATGGATGGAACCTCGAGCTTTATTACCAG 247  
 DB 61 GluileHisapGlyMetAsnLeuLeuLeuTyTyTyGln 73  
 RESULT 12  
 UBL5 CAEEL  
 ID UBL5 CAEEL STANDARD; PRT; 73 AA.  
 AC P91302;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Ubiquitin-like protein 5.  
 GN Name=ubl-5; ORFNames=F46F11.4;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RT "Genome sequencing consortium;  
 RT investigating biology."  
 RL Science 282:2012-2018 (1998).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

-----  
DR EMBL; U88173; AAK21382.1; -; Genomic\_DNA.  
DR PIR; T25763; T25763.  
DR SMR; P91302; 1-73.  
DR Ensemble; F46F11.4; Caenorhabditis elegans.  
DR WormBase; WBGene0006726; ubi-5.  
DR WormPro; F46F11.4; CE10602.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin.1.  
DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
KW Complete proteome; Ubi conjugation pathway.  
FT DOMAIN 1 73 Ubiquitin-like.  
SQ SEQUENCE 73 AA; 8738 MW; 61CA839BBA4006A4 CRC64;

Alignment Scores:  
Pred. No.: 3.42e-30 Length: 73  
Score: 331.00 Matches: 59  
Percent Similarity: 89.0% Conservative: 6  
Best Local Similarity: 80.8% Mismatches: 8  
Query Match: 49.8% Indels: 0  
DB: 1 Gaps: 0

US-10-067-832D-1 (1-370) x UBL5\_CAEEL (1-73)

QY 29 ATGATTGAGTGGTTTGCACACCGCTAGGAAAGAAAGTCCGCTTAAGTGCACACC 88  
DB 1 MetileGlulleThrValAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20  
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGCACTCGTTGGAAT 148  
DB 21 SerAspThrIleGlyAspLeuLysLeuLeuAlaAlaGlnThrGlyThrArgTyrGlu 40  
QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
DB 41 LysileValLeuLysLysLysTyrThrIleTyrLysAspHisileThrLeuMetAspTyr 60  
QY 209 GAAATCCACCATGGGATGAACTGGAGCTTTATTACCAG 247  
DB 61 GluileHisGluGlyPheAsnPheGluLeuTyrTyrGln 73

RESULT 13

Q617M9 CAEBR  
ID Q617M9 CAEBR PRELIMINARY; PRT; 73 AA.  
AC Q617M9  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein CBG14905.  
GN Name=CBG14905;  
OS Caenorhabditis briggsae.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6238;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RG The C. briggsae Sequencing Consortium;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CAAC01000068; CAE68926.1; -; Genomic\_DNA.  
DR SMR; Q617M9; 1-73.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin.1.  
DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 73 AA; 8738 MW; 61CA839BBA4006A4 CRC64;

Alignment Scores:

Pred. No.: 3.42e-30 Length: 73  
Score: 331.00 Matches: 59  
Percent Similarity: 89.0% Conservative: 6  
Best Local Similarity: 80.8% Mismatches: 8  
Query Match: 49.8% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x Q617M9 CAEBR (1-73)

QY 29 ATGATTGAGTGGTTTGCACACCGCTAGGAAAGAAAGTCCGCTTAAGTGCACACC 88  
DB 1 MetileGlulleThrValAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20  
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGCACTCGTTGGAAT 148  
DB 21 SerAspThrIleGlyAspLeuLysLysLeuLeuAlaAlaGlnThrGlyThrArgTyrGlu 40  
QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
DB 41 LysileValLeuLysLysLysTyrThrIleTyrLysAspHisileThrLeuMetAspTyr 60  
QY 209 GAAATCCACCATGGGATGAACTGGAGCTTTATTACCAG 247  
DB 61 GluileHisGluGlyPheAsnPheGluLeuTyrTyrGln 73

RESULT 14

Q5BS57 SCHJA  
ID Q5BS57 SCHJA PRELIMINARY; PRT; 73 AA.  
AC Q5BS57;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Hypothetical protein.  
OS Schistosoma japonicum (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
OC Schistosomatoida; Schistosomatidae; Schistosoma.  
OX NCBI\_TaxID=6182;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Han Z.;  
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY915407; AAX30628.1; -; mRNA.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin.1.  
DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 73 AA; 8679 MW; D3A86229A087CB5 CRC64;

Alignment Scores:

Pred. No.: 5.93e-30 Length: 73  
Score: 329.00 Matches: 58  
Percent Similarity: 89.0% Conservative: 7  
Best Local Similarity: 79.5% Mismatches: 8  
Query Match: 49.5% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x Q5BS57 SCHJA (1-73)

QY 29 ATGATTGAGTGGTTTGCACACCGCTAGGAAAGAAAGTCCGCTTAAGTGCACACC 88  
DB 1 MetileGlulleThrCysAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20  
QY 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAAACTGCACTCGTTGGAAT 148  
DB 21 ThrAspLysValSerAspLeuLysLysLeuLeuAlaAlaGlnThrGlyThrAsnTyrGlu 40  
QY 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
DB 41 ArgileValLeuLysLysLysTyrThrIlePheLysAspHisileThrLeuGlnAspTyr 60  
QY 209 GAAATCCACCATGGGATGAACTGGAGCTTTATTACCAG 247  
DB 61 GluileAsnAspGlyMetAsnLeuGluLeuTyrTyrGln 73

Search completed: April 7, 2006, 23:46:28  
Job time : 171.89 secs

```
RESULT 15
Q4VAF8_MOUSE
ID Q4VAF8_MOUSE PRELIMINARY; PRT; 73 AA.
AC Q4VAF8;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the ENBL/GenBank/DBJ databases.
DR EMBL; BC096400; AAH96400.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8582 MW; 9B2C9977752A6068 CRC64;

Alignment Scores:
Pred. No.: 1.78e-29 Length: 73
Score: 325.00 Matches: 61
Percent Similarity: 86.3% Conservative: 2
Best Local Similarity: 83.6% Mismatches: 10
Query Match: 48.9% Indels: 0
DB: Gaps: 2

US-10-067-832D-1 (1-370) x Q4VAF8_MOUSE (1-73)
Qy 29 ATGATTGAGTGGTGTTCGACGACCGCTCAGGAAAGAAAGTCCGGTTAAGTCAACACC 88
Db 1 MetThrGluValValCysAsnTyrHisLeuGlyLysLysValHisValLysTyrAsnThr 20
Qy 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCAGCTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleGluAlaGlnThrGlyThrHisTyrAsn 40
Qy 149 AAGATCGTCTTAAAGTGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysIleIleLeuLysLysTyrThrPheLeuGlnAspHisValSerLeuGlyAspTyr 60
Qy 209 GAAATCCACCATGGGATGAACCTGGAGCTTTATTATACCAG 247
Db 61 GluIleHisAspGlyMetAsnLeuGluLeuSerTyrGln 73
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:46:49 ; Search time 7.83898 Seconds  
(without alignments)  
780.458 Million cell updates/sec

Title: US-10-067-832D-1  
Perfect score: 665  
Sequence: 1 gtccagagattacagctc.....tttgatgctgcacaaaaaaa 370

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model  
-Q=/abs/ABSSWEB.spool/US10067832/runat.07042006.175615.9797/app.query.fasta\_1  
-DB=Issued Patents AA -OPMT=faстан -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abss802p -USER=US10067832 @CGN 1.1.101 @runat.07042006.175615.9797  
-NCPU=6 -ICPU=3 -NO MWAP -NEG SCORE=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/PCFUS COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/RB COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	394	59.2	73	2	US-09-331-930A-2
2	394	59.2	73	2	US-09-331-930A-19
3	394	59.2	73	2	US-09-331-930A-20
4	394	59.2	73	2	US-09-331-930A-21
5	370	55.6	73	2	US-09-331-930A-24
6	354	53.2	73	2	US-09-331-930A-25
7	331	49.8	73	2	US-09-331-930A-22
8	286	43.0	66	2	US-09-331-930A-23
9	252	37.9	75	2	US-09-248-796A-16302
10	182	27.4	33	2	US-09-331-930A-16
11	182	27.4	33	2	US-09-331-930A-18
12	163	24.5	33	2	US-09-331-930A-14

13	80	12.0	60	2	US-09-331-930A-27
14	73.5	11.1	1078	2	US-09-583-110-4036
15	73.5	11.1	1080	2	US-09-107-433-4843
16	71	10.7	860	2	US-10-290-579A-252
17	69	10.4	184	2	US-09-583-110-4505
18	69	10.4	515	2	US-09-107-532A-6214
19	68	10.2	168	2	US-09-252-991A-21174
20	67.5	10.2	842	2	US-10-290-579A-254
21	67	10.1	512	2	US-09-543-681A-4510
22	66.5	10.0	606	1	US-08-392-806A-6
23	66.5	10.0	606	2	US-09-257-490-6
24	66	9.9	110	2	US-08-252-991A-25531
25	66	9.9	653	2	US-08-849-602C-27
26	66	9.9	653	6	5248599-2
27	66	9.9	921	2	US-09-902-540-15467
28	65.5	9.8	102	2	US-09-732-210-1289
29	65.5	9.8	312	2	US-09-134-001C-3534
30	65.5	9.8	312	2	US-09-710-279-3104
31	65.5	9.8	502	2	US-09-540-236-1922
32	65.5	9.8	637	2	US-09-949-016-10426
33	65.5	9.8	1186	2	US-09-543-681A-5514
34	65	9.8	365	2	US-09-270-767-42958
35	65	9.8	533	2	US-09-252-991A-21407
36	64.5	9.7	281	2	US-09-248-796A-20512
37	64	9.6	214	1	US-08-318-492-4
38	64	9.6	214	1	US-08-707-340-4
39	64	9.6	214	1	US-08-916-902A-3
40	64	9.6	214	1	US-08-994-578-4
41	64	9.6	214	1	US-09-213-389-3
42	64	9.6	782	2	US-09-252-991A-30464
43	63.5	9.5	359	2	US-09-107-532A-6455
44	63.5	9.5	467	2	US-10-040-802-8
45	63.5	9.5	533	2	US-09-107-532A-7241

ALIGNMENTS

RESULT 1  
US-09-331-930A-2  
; Sequence 2, Application US/09331930A  
; Patent No. 6436670  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL Z.  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCES: 22975-20007.00  
; CURRENT APPLICATION NUMBER: US/09/331,930A  
; CURRENT FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP0117/97  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP0323/97  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Amino acid  
; OTHER INFORMATION: sequence for beacon from unknown organism  
US-09-331-930A-2

Alignment Scores:			
Pred. No.:	2.48e-45	Length:	73
Score:	394.00	Matches:	73
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	59.2%	Indels:	0
DB:	2	Gaps:	0

```
US-10-067-832D-1 (1-370) x US-09-331-930A-2 (1-73)
QY 29 ATGATTGAGTGGTTTCACAGCCGCTTAGGAAAGAAAGTCCGCGTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrileGlyAspLeuLysLeuilealaalaglnthrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisaspGlyMetAsnLeuGluLeuTyrTyrGln 73
RESULT 2
US-09-331-930A-19
; Sequence 19, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid
; OTHER INFORMATION: sequence for beacon from unknown organism
US-09-331-930A-19
Alignment Scores:
Pred. No.: 2,48e-45 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 2 Gaps: 0
US-10-067-832D-1 (1-370) x US-09-331-930A-19 (1-73)
QY 29 ATGATTGAGTGGTTTCACAGCCGCTTAGGAAAGAAAGTCCGCGTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrileGlyAspLeuLysLeuilealaalaglnthrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisaspGlyMetAsnLeuGluLeuTyrTyrGln 73
RESULT 3
US-09-331-930A-19
; Sequence 19, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid
; OTHER INFORMATION: sequence for beacon from unknown organism
US-09-331-930A-19
Alignment Scores:
Pred. No.: 2,48e-45 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 2 Gaps: 0
US-10-067-832D-1 (1-370) x US-09-331-930A-19 (1-73)
QY 29 ATGATTGAGTGGTTTCACAGCCGCTTAGGAAAGAAAGTCCGCGTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrileGlyAspLeuLysLeuilealaalaglnthrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisaspGlyMetAsnLeuGluLeuTyrTyrGln 73
RESULT 3
```

```
US-09-331-930A-20
; Sequence 20, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-331-930A-20
Alignment Scores:
Pred. No.: 2,48e-45 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 2 Gaps: 0
US-10-067-832D-1 (1-370) x US-09-331-930A-20 (1-73)
QY 29 ATGATTGAGTGGTTTCACAGCCGCTTAGGAAAGAAAGTCCGCGTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrileGlyAspLeuLysLysLeuilealaalaglnthrGlyThrArgTrpAsn 40
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
QY 209 GAAATCCACGATGGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisaspGlyMetAsnLeuGluLeuTyrTyrGln 73
RESULT 4
US-09-331-930A-21
; Sequence 21, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Murine sp.
```



## US-09-331-930A-21

## Alignment Scores:

Pred. No.: 2,48e-45 Length: 73  
Score: 394.00 Matches: 73  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 59.2% Indels: 0  
DB: 2 Gaps: 0

## US-10-067-832D-1 (1-370) x US-09-331-930A-21 (1-73)

QY 29 ATGATTGAGTGGTTTCACACGCGCTAGGAAGAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAAT 148  
Db 21 AspAspThrileGlyAspLeuLysLysLeuilealaalaglnThrGlyThrArgTrpAsn 40  
QY 149 AAGATCGTTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
Db 41 LysileValleuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTy 60  
QY 209 GAAATCCACGATGGATGAACCTGAGCTTTATTACCG 247  
Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyTrpGln 73

## RESULT 5

## US-09-331-930A-24

; Sequence 24, Application US/09331930A  
; Patent No. 6436670  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL Z.  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975-20007.00  
; CURRENT APPLICATION NUMBER: US/09/331,930A  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP0117/97  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP0323/97  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Oryza sativa

## US-09-331-930A-24

## Alignment Scores:

Pred. No.: 4.95e-42 Length: 73  
Score: 370.00 Matches: 67  
Percent Similarity: 97.2% Conservative: 3  
Best Local Similarity: 93.1% Mismatches: 2  
Query Match: 55.6% Indels: 0  
DB: 2 Gaps: 0

## US-10-067-832D-1 (1-370) x US-09-331-930A-24 (1-73)

QY 29 ATGATTGAGTGGTTTCACACGCGCTAGGAAGAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAAT 148  
Db 21 AspAspThrileGlyAspLeuLysLysLeuilealaalaglnThrGlyThrArgTrpAsn 40  
QY 149 AAGATCGTTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208

Db 41 LysileValleuLysLysTrpTyThrileTyLysAspHisileThrLeuAlaAspTy 60  
QY 209 GAAATCCACGATGGATGAACCTGAGCTTTATTAC 244  
Db 61 GluileHisAspGlyMetGlyLeuGluLeuTyTr 72

## RESULT 6

US-09-331-930A-25  
; Sequence 25, Application US/09331930A  
; Patent No. 6436670

; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL Z.  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975-20007.00  
; CURRENT APPLICATION NUMBER: US/09/331,930A  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP0117/97  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP0323/97  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae

## US-09-331-930A-25

## Alignment Scores:

Pred. No.: 7.86e-40 Length: 73  
Score: 354.00 Matches: 65  
Percent Similarity: 93.1% Conservative: 2  
Best Local Similarity: 50.3% Mismatches: 5  
Query Match: 53.2% Indels: 0  
DB: 2 Gaps: 0

## US-10-067-832D-1 (1-370) x US-09-331-930A-25 (1-73)

QY 29 ATGATTGAGTGGTTTCACACGCGCTAGGAAGAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAAT 148  
Db 21 AspAspThrileGlyAspLeuLysLysLeuilealaalaglnThrGlyThrArgTrpAsn 40  
QY 149 AAGATCGTTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
Db 41 LysileValleuLysLysTrpTyThrileTyLysAspHisileThrLeuAlaAspTy 60  
QY 209 GAAATCCACGATGGATGAACCTGAGCTTTATTAC 244  
Db 61 GluValHisAspGlnThrAsnLeuGluLeuTyTr 72

## RESULT 7

US-09-331-930A-22  
; Sequence 22, Application US/09331930A  
; Patent No. 6436670

; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL Z.  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975-20007.00  
; CURRENT APPLICATION NUMBER: US/09/331,930A  
; PRIOR FILING DATE: 1999-06-30  
; CURRENT APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP0117/97  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP0323/97

```
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 22
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-331-930A-22

Alignment Scores:
Pred. No.: 1.15e-36 Length: 73
Score: 331.00 Matches: 59
Percent Similarity: 89.0% Conservative: 6
Best Local Similarity: 80.8% Mismatches: 8
Query Match: 49.8% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-331-930A-22 (1-73)

QY 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAGAAAGTCCGCTTAAGTGAACACC 88
Db 1 MetileGluThrValAsnAspArgLeuGlyLysValArgLysCysAsnPro 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 SerAspThrIleGlyAspLeuLysLysLysLysLysLysLysLysLysLysLysLys 40
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValLeuLysLysTptYrThrIleTyLysAspHisileThrLeuMetAspTyr 60
QY 209 GAAATCCAGATCGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisGlyPheAsnPheGluLeuTyTyGln 73

RESULT 8
US-09-331-930A-23
; Sequence 23, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR FILING DATE: 1997-11-11
; PRIOR FILING DATE: 1997-10-31
; PRIOR FILING DATE: 1997-10-31
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 23
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Fasiola hepatica
US-09-331-930A-23

Alignment Scores:
Pred. No.: 1.7e-30 Length: 66
Score: 286.00 Matches: 54
Percent Similarity: 89.4% Conservative: 5
Best Local Similarity: 81.8% Mismatches: 7
Query Match: 43.0% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-331-930A-23 (1-66)

QY 50 GACCGTCTAGGAAGAAAGTCCGCTTAAGTGAACACCGATGACACCATCGGGGACTTG 109
Db 1 AspArgLeuGlyLysValArgValLysCysAsnProThrAspLysValGlyAspLeu 20
```

```
QY 110 AAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAATAAGATCGTCTTAAAGAGTGG 169
Db 21 LysLysLeuIleAlaAlaGlnThrGlyThrAlaProGluArgIleValLeuLysLysTyr 40
QY 170 TACACGATTTTAAAGGACCATGTATCTCTGGGAGATTATGAAATCCACGATGGATGAAC 229
Db 41 TyrThrIleTyLysAspHisValThrLeuArgAspTyrgluIleAsnAspGlyMetAsn 60
QY 230 CTGGAGCTTTATTACCAG 247
Db 61 LeuGluLeuTyTyGln 66

RESULT 9
US-09-248-796A-16302
; Sequence 16302, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16302
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16302

Alignment Scores:
Pred. No.: 8.5e-26 Length: 75
Score: 252.00 Matches: 44
Percent Similarity: 77.8% Conservative: 12
Best Local Similarity: 61.1% Mismatches: 16
Query Match: 37.9% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-248-796A-16302 (1-75)

QY 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAGAAAGTCCGCTTAAGTGAACACC 88
Db 3 MetileGluThrValAsnAspArgLeuGlyLysLysLysLysLysLysLysLysLysLys 22
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148
Db 23 ThrAspThrIleGlyAspValLysLysLysLysLysLysLysLysLysLysLysLys 42
QY 149 AAGATCGTTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 43 LysileLeuLysLysGlyTyrglnValTyLysAspHisileThrLeuAspAspTyr 62
QY 209 GAAATCCAGATGGATGAACCTGGAGCTTTATTACC 244
Db 63 GluileHisAspGlyPheAsnPheGluLeuTyTyGln 74

RESULT 10
US-09-331-930A-16
; Sequence 16, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
```

```
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Psammomys obesus
US-09-331-930A-16

Alignment Scores:
Pred. No.: 2,61e-16 Length: 33
Score: 182.00 Matches: 33
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 27.4% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-331-930A-16 (1-33)

QY 149 AAGATCGTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGAGATTAT 208
Db 1 LysileValLeuLysIleThrIlePheLysAspHisValSerLeuGlyAspTyr 20
QY 209 GAAATCCAGCATGGATGAACCTGGAGCTTTATTACCAG 247
Db 21 GluIleHisAspGlyMetAsnLeuGluLeuTyrTyrGln 33

RESULT 11
US-09-331-930A-18
; Sequence 18, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 18
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Psammomys obesus
US-09-331-930A-18

Alignment Scores:
Pred. No.: 2,61e-16 Length: 33
Score: 182.00 Matches: 33
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 27.4% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-331-930A-18 (1-33)

QY 149 AAGATCGTCTTAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGAGATTAT 208
Db 1 LysileValLeuLysIleThrIlePheLysAspHisValSerLeuGlyAspTyr 20
QY 209 GAAATCCAGCATGGATGAACCTGGAGCTTTATTACCAG 247
Db 21 GluIleHisAspGlyMetAsnLeuGluLeuTyrTyrGln 33
```

```
RESULT 12
US-09-331-930A-14
; Sequence 14, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (15)
; OTHER INFORMATION: variable amino acid
US-09-331-930A-14

Alignment Scores:
Pred. No.: 1,07e-13 Length: 33
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 24.5% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-1 (1-370) x US-09-331-930A-14 (1-33)

QY 29 ATGATTGAGTGTGTTTGCACAGCCGCTAGGAAGAACTCCGCTTAGTGCACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysVal***ValLysCysAsnThr 20
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCC 127
Db 21 AspAspThrileGlyAspLeuLysLeuIleAlaIala 33

RESULT 13
US-09-331-930A-27
; Sequence 27, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-331-930A-27
```



Qy 227 AACCTGGAGCTTTATTACCAG 247  
Db 199 LysLeuGluLeuPheTyrGlu 205

Search completed: April 7, 2006, 23:49:33  
Job time : 41.1949 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:48:08 ; Search time 26.8877 Seconds  
(without alignments)  
1149.944 Million cell updates/sec

Title: US-10-067-832D-1  
Perfect score: 665  
Sequence: 1 gttccagagattacagctc.....tttgatgtctgcaaaaaaaa 370

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat\_07042006\_175617\_9835/app\_query.fasta\_1  
-DB=Published Applications AA\_Main -QPM=fastan -SUFFIX=n2p.rapbm  
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USER=US10067832 @CGN 1.1 405 @runat\_07042006\_175617\_9835 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418	62.9	115	4	US-10-264-049-2463 Sequence 2463, Ap
2	394	59.2	73	5	US-10-067-832D-2 Sequence 2, Appl
3	394	59.2	73	5	US-10-067-832D-15 Sequence 15, Appl
4	394	59.2	73	5	US-10-067-832D-16 Sequence 16, Appl
5	394	59.2	73	5	US-10-067-832D-68 Sequence 68, Appl
6	370	55.6	73	5	US-10-067-832D-19 Sequence 19, Appl
7	366	55.0	68	4	US-10-424-599-250542 Sequence 250542,
8	354	53.2	73	5	US-10-067-832D-20 Sequence 20, Appl
9	339	51.0	73	6	US-11-097-143-6522 Sequence 6522, Ap
10	331	49.8	73	5	US-10-067-832D-17 Sequence 17, Appl
11	319	48.0	73	4	US-10-424-599-212296 Sequence 212296,

12	312	46.9	73	4	US-10-437-963-147593 Sequence 147593,
13	312	46.9	73	4	US-10-767-701-46049 Sequence 46049, A
14	312	46.9	73	4	US-10-425-115-233843 Sequence 233843,
15	312	46.9	73	4	US-10-425-115-233846 Sequence 233846,
16	312	46.9	73	4	US-10-425-115-233848 Sequence 233848,
17	312	46.9	73	4	US-10-425-115-233849 Sequence 233849,
18	312	46.9	73	4	US-10-425-115-233851 Sequence 233851,
19	312	46.9	73	4	US-10-425-115-233853 Sequence 233853,
20	311	46.8	73	4	US-10-425-115-288480 Sequence 288480,
21	311	46.8	73	4	US-10-425-115-288482 Sequence 288482,
22	311	46.8	73	4	US-10-425-115-288483 Sequence 288483,
23	311	46.8	73	4	US-10-425-115-288486 Sequence 288486,
24	304.5	45.8	74	4	US-10-424-599-207467 Sequence 207467,
25	300	45.1	73	4	US-10-425-115-288481 Sequence 288481,
26	286	43.0	66	5	US-10-067-832D-18 Sequence 18, Appl
27	285	42.9	73	5	US-10-739-930-8461 Sequence 8461, Ap
28	281	42.3	73	4	US-10-425-115-288479 Sequence 288479,
29	252	37.9	89	4	US-10-083-357-770 Sequence 770, App
30	252	37.9	320	4	US-10-338-411-11 Sequence 11, Appl
31	252	37.9	320	4	US-10-389-640-11 Sequence 11, Appl
32	249	37.4	89	4	US-10-425-115-355452 Sequence 355452,
33	248	37.3	73	4	US-10-425-115-290033 Sequence 290033,
34	244	36.7	82	4	US-10-424-599-191611 Sequence 191611,
35	234	35.2	71	4	US-10-424-599-147197 Sequence 147197,
36	194	29.2	110	4	US-10-424-599-213543 Sequence 213543,
37	184	27.7	70	4	US-10-425-115-302525 Sequence 302525,
38	179.5	27.0	79	4	US-10-425-115-269824 Sequence 269824,
39	170	25.6	63	4	US-10-424-599-229016 Sequence 229016,
40	163	24.5	33	5	US-10-067-832D-14 Sequence 14, Appl
41	160.5	24.1	93	4	US-10-424-599-266457 Sequence 266457,
42	160.5	24.1	112	4	US-10-425-114-47750 Sequence 47750, A
43	160.5	24.1	112	4	US-10-425-114-59753 Sequence 59753, A
44	158	23.8	112	4	US-10-425-114-62746 Sequence 62746, A
45	158	23.8	113	4	US-10-425-114-61768 Sequence 61768, A

#### ALIGNMENTS

##### RESULT 1

US-10-264-049-2463  
; Sequence 2463, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birste et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P4133p1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn ver. 3.1  
; SEQ ID NO 2463  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-2463

Alignment Scores:  
Pred. No.: 8,74e-44 Length: 115  
Score: 418.00 Matches: 78  
Percent Similarity: 96.3% Conservative: 1  
Best Local Similarity: 95.1% Mismatches: 3  
Query Match: 62.9% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-1 (1-370) x US-10-264-049-2463 (1-115)

QY 2 TTCAGAGATTACGCTCCAGCCACATGATTCAGTGGTTTCACGACCGCTAGGA 61  
|||||  
Db 34 PheGluAlaIleArgAlaProIleGluValValCysAsnArgArgLeuGly 53  
|||||

QY 62 AAGAAAGTCCGCTTAAGTCAACACCGATGACACCATCGGGACTTGAAGAACTGATA 121  
Db |||||||  
QY 54 LysLysValArgValLysCysAsnThrAspAspThrIleGlyAspLeuLysLysLeuLe 73  
Db |||||||  
QY 122 GCGGCCCAAACTGGCACTCGTGGAAATAGATCTTCTTAAAGTGGTACACGATTTT 181  
Db |||||||  
QY 74 AlaAlaGlnThrGlyThrArgTrpAsnLysIleValLeuLysLysTrpThrIlePhe 93  
Db |||||||  
QY 182 AAGGACCATATCTCTGGGAGATTATGAATCCACGATCGGATGAACCTGGAGCTTAT 241  
Db |||||||  
QY 94 LysAspHisValSerLeuGlyAspTyrGluIleHisAspGlyMetAsnLeuGluLeuTyr 113  
Db |||||||  
QY 242 TACCAG 247  
Db |||||||  
QY 114 TyrGln 115

## RESULT 2

US-10-067-832D-2  
; Sequence 2, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 2  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Psammomys obesus  
US-10-067-832D-2

Alignment Scores:  
Pred. No.: 8.76e-41 Length: 73  
Score: 394.00 Matches: 73  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 59.2% Indels: 0  
DB: 5 Gaps: 0

## US-10-067-832D-1 (1-370) x US-10-067-832D-2 (1-73)

QY 29 ATGATTGAGTGGTTTTCACAGCCGCTAGGAAGAAAGTCCGCGTTAAGTCAACACC 88  
Db 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40  
QY 149 AAGATCGTTCTTAAAAAGTGGTACACCATGTTTAAAGACCATGTATCTCTGGGAGATTAT 208  
Db |||||||  
QY 41 LysIleValLeuLysLysTrpThrIlePheLysAspHisValSerLeuGlyAspTyr 60  
QY 209 GAAATCCACGATGGATGAACCTGGAGCTTTATTACCAG 247  
Db |||||||  
QY 61 GluIleHisAspGlyMetAsnLeuGluLeuTyrTrpGln 73

## RESULT 3

US-10-067-832D-15  
; Sequence 15, Application US/10067832D

Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 15  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-067-832D-15

Alignment Scores:  
Pred. No.: 8.76e-41 Length: 73  
Score: 394.00 Matches: 73  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 59.2% Indels: 0  
DB: 5 Gaps: 0

## US-10-067-832D-1 (1-370) x US-10-067-832D-15 (1-73)

QY 29 ATGATTGAGTGGTTTTCACAGCCGCTAGGAAGAAAGTCCGCGTTAAGTCAACACC 88  
Db 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
QY 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40  
QY 149 AAGATCGTTCTTAAAAAGTGGTACACCATGTTTAAAGACCATGTATCTCTGGGAGATTAT 208  
Db |||||||  
QY 41 LysIleValLeuLysLysTrpThrIlePheLysAspHisValSerLeuGlyAspTyr 60  
QY 209 GAAATCCACGATGGATGAACCTGGAGCTTTATTACCAG 247  
Db |||||||  
QY 61 GluIleHisAspGlyMetAsnLeuGluLeuTyrTrpGln 73  
RESULT 4  
US-10-067-832D-16  
; Sequence 16, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 16  
; LENGTH: 73



```
;
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-067-832D-16

Alignment Scores:
Pred. No.: 8,76e-41 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 5 Gaps: 0

US-10-067-832D-1 (1-370) x US-10-067-832D-16 (1-73)

Qy 29 ATGATTGAGTGGTTCGACACCGTCTAGGAAAGAAAGTCCGCTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
Qy 89 GATCACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40
Qy 149 AAGATCGTTCCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTCGGAGATTAT 208
Db 41 LysileValLeuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
Qy 209 GAAATCACGATGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyrTyrGln 73

RESULT 5
US-10-965-898-68
; Sequence 68, Application US/10965898
; Publication No. US20050084936A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Bandman, Olga
; Hillman, Jennifer L.
; Au-Young, Janice
; Tang, Y. Tom
; Yue, Henry
; Shah, Purvi
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
; NUMBER OF SEQUENCES: 150
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/965,898
; FILING DATE: 18-Oct-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,403
; FILING DATE: 31-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0455 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 68:

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT19
; CLONE: 3245954
; SEQUENCE DESCRIPTION: SEQ ID NO: 68 :
US-10-965-898-68

Alignment Scores:
Pred. No.: 8,76e-41 Length: 73
Score: 394.00 Matches: 73
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 5 Gaps: 0

US-10-067-832D-1 (1-370) x US-10-965-898-68 (1-73)

Qy 29 ATGATTGAGTGGTTCGACACCGTCTAGGAAAGAAAGTCCGCTTAAGTCAACACC 88
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
Qy 89 GATCACACCATCGGGGACTTGAAGAACTGATAGCGCCCAAACTGGCACTCGTTGGAAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaGlnThrGlyThrArgTrpAsn 40
Qy 149 AAGATCGTTCCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTCGGAGATTAT 208
Db 41 LysileValLeuLysLysTrpTyThrilePheLysAspHisValSerLeuGlyAspTyr 60
Qy 209 GAAATCACGATGGATGAACCTGGAGCTTTATTACCAG 247
Db 61 GluileHisAspGlyMetAsnLeuGluLeuTyrTyrGln 73

RESULT 6
US-10-067-832D-19
; Sequence 19, Application US/10067832D
; Publication No. US20050059108A1
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL ZEV
; COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975200701
; CURRENT APPLICATION NUMBER: US/10/067,832D
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 09/331,930
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP 0117
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP 0323
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 19
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-067-832D-19

Alignment Scores:
Pred. No.: 9,87e-38 Length: 73
Score: 370.00 Matches: 67
Percent Similarity: 97.2% Conservative: 3
Best Local Similarity: 93.1% Mismatches: 2
Query Match: 55.6% Indels: 0
DB: 5 Gaps: 0

US-10-067-832D-1 (1-370) x US-10-067-832D-19 (1-73)
```

```
QY 29 ATGATTGAGTGGTTCGACACGCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
Db 1 MetilegluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGACTTGAAGAAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTTrpAsn 40
QY 149 AGATCGTTCCTTAAAAAGTGGTACAGATTTTAAAGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValLeuLysLysTTrpThrIleTyrLysAspHisIleThrLeuAlaAspTyr 60
QY 209 GAAATCCACCATGGGATGACCTGGAGCTTTATTAC 244
Db 61 GluileHisaspGlyMetGlyLeuGluLeuTyrTyr 72
RESULT 7
US-10-424-599-250542
; Sequence 250542, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250542
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6826C.1.pep
US-10-424-599-250542
Alignment Scores:
Pred. No.: 3,13e-37 Length: 68
Score: 366.00 Matches: 68
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 55.0% Indels: 0
Gaps: 4
US-10-067-832D-1 (1-370) x US-10-424-599-250542 (1-68)
QY 29 ATGATTGAGTGGTTCGACACGCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
Db 1 MetilegluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGACTTGAAGAAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTTrpAsn 40
QY 149 AGATCGTTCCTTAAAAAGTGGTACAGATTTTAAAGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValLeuLysLysTTrpThrIleTyrLysAspHisIleThrLeuAlaAspTyr 60
QY 209 GAAATCCACCATGGGATGACCTGGAGCTTTATTAC 232
Db 61 GluileHisaspGlyMetAsnLeu 68
RESULT 8
US-10-067-832D-20
; Sequence 20, Application US/10067832D
; Publication No. US20050059108A1
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL ZEV
; APPLICANT: COLLIER, GREGORY
```

```
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 229752000701
; CURRENT APPLICATION NUMBER: US/10/067,832D
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 09/331,930
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP 0117
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP 0323
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-067-832D-20
Alignment Scores:
Pred. No.: 1,07e-35 Length: 73
Score: 354.00 Matches: 65
Percent Similarity: 93.1% Conservative: 2
Best Local Similarity: 90.3% Mismatches: 5
Query Match: 53.2% Indels: 0
Gaps: 0
US-10-067-832D-1 (1-370) x US-10-067-832D-20 (1-73)
QY 29 ATGATTGAGTGGTTCGACACGCGTCTAGGAAGAAAGTCCGCTTAAGTGCACACC 88
Db 1 MetilegluValValCysAsnAspArgLeuGlyLysValArgValLysCysAsnThr 20
QY 89 GATGACACCATCGGGACTTGAAGAAACTGATAGCGGCCCAAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaGlnThrGlyThrArgTTrpAsn 40
QY 149 AGATCGTTCCTTAAAAAGTGGTACAGATTTTAAAGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValLeuLysLysTTrpThrIleTyrLysAspHisIleThrLeuAlaAspTyr 60
QY 209 GAAATCCACCATGGGATGACCTGGAGCTTTATTAC 244
Db 61 GluValHisaspGlnThrAsnLeuGluLeuTyrTyr 72
RESULT 9
US-11-097-143-6522
; Sequence 6522, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6522
; LENGTH: 73
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-6522

Alignment Scores:
Pred. No.:      8.64e-34      Length:      73
Score:          339.00      Matches:      63
Percent Similarity: 90.4%      Conservative: 3
Best Local Similarity: 86.3%      Mismatches: 7
Query Match:      51.0%      Indels:      0
DB:              6          Gaps:      0

US-10-067-832D-1 (1-370) x US-11-097-143-6522 (1-73)

Qy 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAGAAAGTCCCGGTTAAGTGCACACC 88
Db 1 MetileGluileThrValAsnAspArgLeuGlyLysValArgValLysCysAsnPro 20
Qy 89 GATCACACCATCGGGGACTTGAAGAACTGATAGCGCCCAACTGGCACTCGTTGGAAT 148
Db 21 SerAspThrIleGlyAspLeuLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Qy 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrIleTyLysAspHisIleThrLeuMetAspTy 60
Qy 209 GAAATCCACGATGGATGAACCTCGAGCTTTATTACCAG 247
Db 61 GluileHisGlyMetAsnPheGluLeuTyTrpGln 73

RESULT 10
US-10-067-832D-17
; Sequence 17, Application US/10067832D
; Publication No. US20050059108A1
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL ZEV
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 229752000701
; CURRENT APPLICATION NUMBER: US/10/067,832D
; PRIOR FILING DATE: 2002-06-03
; PRIOR FILING DATE: 1999-06-30
; PRIOR FILING DATE: 1998-10-30
; PRIOR FILING DATE: 1997-10-31
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 17
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-067-832D-17

Alignment Scores:
Pred. No.:      9e-33      Length:      73
Score:          331.00      Matches:      59
Percent Similarity: 89.0%      Conservative: 6
Best Local Similarity: 80.8%      Mismatches: 8
Query Match:      49.8%      Indels:      0
DB:              5          Gaps:      0

US-10-067-832D-1 (1-370) x US-10-067-832D-17 (1-73)
```

```
Qy 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAGAAAGTCCCGGTTAAGTGCACACC 88
Db 1 MetileGluileThrValAsnAspArgLeuGlyLysValArgValLysCysAsnPro 20
Qy 89 GATCACACCATCGGGGACTTGAAGAACTGATAGCGCCCAACTGGCACTCGTTGGAAT 148
Db 21 SerAspThrIleGlyAspLeuLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Qy 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileValleuLysLysTrpTyThrIleTyLysAspHisIleThrLeuMetAspTy 60
Qy 209 GAAATCCACGATGGATGAACCTCGAGCTTTATTACCAG 247
Db 61 GluileHisGlyMetAsnPheGluLeuTyTrpGln 73

RESULT 11
US-10-424-599-212296
; Sequence 212296, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212296
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33729C.1.pep
US-10-424-599-212296

Alignment Scores:
Pred. No.:      3.02e-31      Length:      73
Score:          319.00      Matches:      59
Percent Similarity: 91.7%      Conservative: 7
Best Local Similarity: 81.9%      Mismatches: 6
Query Match:      48.0%      Indels:      0
DB:              4          Gaps:      0

US-10-067-832D-1 (1-370) x US-10-424-599-212296 (1-73)

Qy 29 ATGATTGAGTGGTTTGCACGACCGTCTAGGAAGAAAGTCCCGGTTAAGTGCACACC 88
Db 1 MetileGluValValLeuAsnAspArgLeuGlyLysValArgValLysCysAsnAsp 20
Qy 89 GATCACACCATCGGGGACTTGAAGAACTGATAGCGCCCAACTGGCACTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLeuValLeuLeuLeuLeuLeuLeuLeuLeu 40
Qy 149 AAGATCGTTCTTAAAAAGTGGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208
Db 41 LysileArgileGlnLysTrpTyThrIleTyLysAspHisIleThrLeuLysAspTy 60
Qy 209 GAAATCCACGATGGATGAACCTCGAGCTTTATTACC 244
Db 61 GluileHisAspGlyMetGlyLeuGluLeuTyTrp 72

RESULT 12
US-10-437-963-147593
; Sequence 147593, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 147593  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48107C.1.pep  
US-10-437-963-147593

Alignment Scores:  
Pred. No.: 2,35e-30 Length: 73  
Score: 312.00 Matches: 58  
Percent Similarity: 88.9% Conservative: 6  
Best Local Similarity: 80.6% Mismatches: 8  
Query Match: 46.9% Indels: 0  
Gaps: 0  
DB:

US-10-067-832D-1 (1-370) x US-10-437-963-147593 (1-73)  
Qy 29 ATGATTGAGTGGTTTGCACGACCGCTTAGGAAAGAAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgValLysCysAsnGlu 20  
Qy 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAACTGCGACTCGTTGGAAT 148  
Db 21 AspAspThrIleGlyAspLeuLysLeuValAlaAlaGlnThrGlyThrArgProGlu 40  
Qy 149 AAGATCGTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
Db 41 LysIleArgIleGlnLysTrpTyrAsnIleTyrLysAspHisIleThrLeuLysAspTyr 60  
Qy 209 GAAATCCAGATGGATGAACCTGGAGCTTTATTAC 244  
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

RESULT 13  
US-10-767-701-46049  
; Sequence 46049, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 46049  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C594\_1.pep  
US-10-767-701-46049

Alignment Scores:  
Pred. No.: 2,35e-30 Length: 73  
Score: 312.00 Matches: 58  
Percent Similarity: 88.9% Conservative: 6  
Best Local Similarity: 80.6% Mismatches: 8  
Query Match: 46.9% Indels: 0  
DB:

DB: 4 Gaps: 0  
US-10-067-832D-1 (1-370) x US-10-767-701-46049 (1-73)  
Qy 29 ATGATTGAGTGGTTTGCACGACCGCTTAGGAAAGAAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgValLysCysAsnGlu 20  
Qy 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAACTGCGACTCGTTGGAAT 148  
Db 21 AspAspThrIleGlyAspLeuLysLeuValAlaAlaGlnThrGlyThrArgProGlu 40  
Qy 149 AAGATCGTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
Db 41 LysIleArgIleGlnLysTrpTyrAsnIleTyrLysAspHisIleThrLeuLysAspTyr 60  
Qy 209 GAAATCCAGATGGATGAACCTGGAGCTTTATTAC 244  
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

RESULT 14  
US-10-425-115-233843  
; Sequence 233843, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 233843  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_144858C.1.pep  
US-10-425-115-233843

Alignment Scores:  
Pred. No.: 2,35e-30 Length: 73  
Score: 312.00 Matches: 58  
Percent Similarity: 88.9% Conservative: 6  
Best Local Similarity: 80.6% Mismatches: 8  
Query Match: 46.9% Indels: 0  
Gaps: 0  
DB:

US-10-067-832D-1 (1-370) x US-10-425-115-233843 (1-73)  
Qy 29 ATGATTGAGTGGTTTGCACGACCGCTTAGGAAAGAAAGTCCGCGTTAAGTGCACACC 88  
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgValLysCysAsnGlu 20  
Qy 89 GATGACACCATCGGGACTTGAAGAACTGATAGCGGCCCAACTGCGACTCGTTGGAAT 148  
Db 21 AspAspThrIleGlyAspLeuLysLeuValAlaAlaGlnThrGlyThrArgProGlu 40  
Qy 149 AAGATCGTCTTAAAAAGTGTACACGATTTTAAAGGACCATGTATCTCTGGGAGATTAT 208  
Db 41 LysIleArgIleGlnLysTrpTyrAsnIleTyrLysAspHisIleThrLeuLysAspTyr 60  
Qy 209 GAAATCCAGATGGATGAACCTGGAGCTTTATTAC 244  
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

RESULT 15  
US-10-425-115-233846  
; Sequence 233846, Application US/10425115  
; Publication No. US20040214272A1

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 233846
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144860C.1.pep
US-10-425-115-233846

```

```

Alignment Scores:
Pred. No.: 2.35e-30 Length: 73
Score: 312.00 Matches: 58
Percent Similarity: 88.9% Conservative: 6
Best Local Similarity: 80.6% Mismatches: 8
Query Match: 46.9% Indels: 0
DB: 4 Gaps: 0

US-10-067-832D-1 (1-370) x US-10-425-115-233846 (1-73)

Qy 29 ATGATTGAGTGGTTGCAACGCGCTTAGAAGAAGTCCGGTTAAGTGCACACC 88
Db 1 MetIleGluValValLeuAsnAspArgLeuGlyLysLysValArgValLysCysAsnGlu 20

Qy 89 GATGACACCATCGGGGACTTGAAGAACTGATAGCGGCCCAAACTGGCAGCTCGTTGGAAT 148
Db 21 AspAspThrIleGlyAspLeuLysLeuValalaGlnThrGlyThrArgProGlu 40

Qy 149 AAGATCGTTCTTAAAGTGTGTACACGATTTTAAAGACCATGTATCTCTCGGAGATTAT 208
Db 41 LysIleArgIleGlnLysTrpTyrAsnIleTyrLysAspHisIleThrLeuLysAspTyr 60

Qy 209 GAATCCACGATGGGATCAACCTGGAGCTTTATTAC 244
Db 61 GluIleHisAspGlyMetGlyLeuGluLeuTyrTyr 72

```

Search completed: April 7, 2006, 23:55:20  
Job time : 136.439 secs

**THIS PAGE BLANK (USPTO)**

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: April 7, 2006, 23:49:48 ; Search time 3.60593 Seconds  
(without alignments)  
640.114 Million cell updates

**Title:** US-10-067-832D-1

Perfect score:

Sequence: 1 gttccaggagattacagctc.....tttgatgctgcaaaaaaaaa 370

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 368322

Minimum DB seq length: 0

Maximum DB seq length: 2000000000  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

**Command line parameters:**

```

MODEL=frame_n2p2.model -DEV=xlp
-Q/-absp/ABSWER spoof/US1067832/runat_07042006_175620_9893/app_query.fasta.1
-DB=published Applications AA NEW -QFMT=fastan -SUFFIX=n2p2.rapbn -MINMATCH=0.1
-LOOCPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATTR=chrom62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=100 -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTTMP=ptc0 -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=as02p
-USER=US1067832 @CGEN 1 1 34 @runat_07042006_175620_9893 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORE=0 -WAIT -DSPBLOC=100 -LONGLOG -DEXTIMEOUT=120
-WARN TIMEOUT=30 -THRTADS=1 -XGAPOP=1 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0 -DELOP=5 -DELOPEXT=7

```

Database : Published Applications AA New:\*

	newsworld_applications_Ar_new*
1:	/SID55/pdodata/1/pubppaa/US08 NEW PUB pep.*
2:	/SID55/pdodata/1/pubppaa/US06 NEW PUB pep.*
3:	/SID55/pdodata/1/pubppaa/US07 NEW PUB pep.*
4:	/SID55/pdodata/1/pubppaa/PCT NEW PUB pep.*
5:	/SID55/pdodata/1/pubppaa/US09 NEW PUB pep.*
6:	/SID55/pdodata/1/pubppaa/US10 NEW PUB pep.*
7:	/SID55/pdodata/1/pubppaa/US11 NEW PUB pep.*
8:	/SID55/pdodata/1/pubppaa/US60 NEW PUB pep.*

4. Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	72.5	10.9	860	7	US-11-100-356-6		Sequence 6, Appli
2	72.5	10.9	867	7	US-11-100-356-3		Sequence 3, Appli
3	69.5	10.5	135	5	US-09-978-360A-460		Sequence 460, App
4	67.5	10.2	869	7	US-11-100-356-4		Sequence 4, Appli
5	66.5	10.0	315	7	US-11-096-588A-33836		Sequence 33836, A
6	66.5	10.0	332	7	US-11-096-568A-33835		Sequence 33835, A
7	66.5	10.0	391	7	US-11-096-568A-33834		Sequence 33834, A
8	66	9.9	495	7	US-11-182-016-31		Sequence 31, Appli
9	65.5	9.8	312	6	US-10-793-626-3104		Sequence 3104, Ap

C	10	65.5	9.8	371	7	US-11-087-099-5646	Sequence 6646, App
	11	65.5	9.8	650	7	US-11-119-249-7	Sequence 7, Appl
	12	65	9.8	654	6	US-10-510-947-6	Sequence 6, Appl
	C 13	64.5	9.7	501	7	US-11-172-740-441	Sequence 441, App
	C 14	64.5	9.7	668	7	US-11-095-568A-31245	Sequence 31245, A
	C 15	64.5	9.7	690	7	US-11-095-568A-31244	Sequence 31244, A
	C 16	64.5	9.7	682	7	US-11-095-568A-31243	Sequence 31243, A
	C 17	63.5	9.5	258	7	US-11-058-924-6	Sequence 6, Appl
	C 18	63	9.5	239	6	US-10-957-569-54	Sequence 54, Appl
	C 19	63	9.5	239	7	US-11-097-589-53	Sequence 53, Appl
	C 20	62.5	9.4	137	5	US-09-978-360A-765	Sequence 765, App
	C 21	62.5	9.4	310	7	US-11-087-099-11274	Sequence 11274, A
	C 22	62.5	9.4	360	7	US-11-087-099-11236	Sequence 11236, A
	C 23	62.5	9.4	412	6	US-10-485-537-204	Sequence 204, App
	C 24	62.5	9.4	463	7	US-11-087-099-9184	Sequence 9184, App
	C 25	61.5	9.2	377	7	US-11-024-959-350	Sequence 350, App
	C 26	61.5	9.2	565	6	US-10-506-454-698	Sequence 698, App
	C 27	61	9.2	150	7	US-11-095-568A-26888	Sequence 26888, A
	C 28	61	9.2	150	7	US-11-095-568A-26896	Sequence 26896, A
	C 29	61	9.2	175	7	US-11-095-568A-26895	Sequence 26895, A
	C 30	61	9.2	177	7	US-11-095-568A-26887	Sequence 26887, A
	C 31	61	9.2	463	7	US-11-072-512-3058	Sequence 3058, App
	C 32	60.5	9.1	392	6	US-10-498-691A-7	Sequence 7, Appl
	C 33	60.5	9.1	403	7	US-11-095-568A-32637	Sequence 32637, A
	C 34	60.5	9.1	426	7	US-11-095-568A-32636	Sequence 32636, A
	C 35	60.5	9.1	566	6	US-10-203-486-13	Sequence 13, Appl
	C 36	60.5	9.1	566	7	US-11-095-568A-32635	Sequence 32635, A
	C 37	60.5	9.1	1273	7	US-11-181-330-8	Sequence 8, Appl
	C 38	60	9.0	150	7	US-11-172-740-146	Sequence 146, App
	C 39	60	9.0	513	7	US-11-098-686-10835	Sequence 10835, A
	C 40	59.5	8.9	1018	7	US-11-095-568A-34025	Sequence 34025, A
	C 41	59.5	8.9	1045	7	US-11-095-568A-34024	Sequence 34024, A
	C 42	59.5	8.9	1111	7	US-11-142-700-26	Sequence 26, Appl
	C 43	59.5	8.9	1111	7	US-11-095-568A-34023	Sequence 34023, A
	C 44	59	8.9	523	6	US-10-467-657-5392	Sequence 5392, App
	C 45	58.5	8.8	371	7	US-11-087-099-5960	Sequence 5960, App

## ALIGNMENTS

```

RESULT 1
US-11-100-356-6
; Sequence 6, Application US/11100356
; Publication No. US20060057115A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, AND USES THEREOF
; FILE REFERENCE: 2300-1621-20
; CURRENT APPLICATION NUMBER: US/11/100,356
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: US/10/190,434
; PRIOR FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: sequence

```

Alignment Scores:	
Pred. No.:	1.41
Score:	72.50
Percent Similarity:	41.1%
Best Local Similarity:	31.0%
Query Match:	10.9%
DB:	7
Matches:	Length:
Conservative:	13
Mismatches:	34
Indels:	42
Gaps:	10
	860

US-10-067-832D-1 (1-370) x US-11-100-356-6 (1-860)

Qy 32 ATTGAGTGGTTTGC-----AACGACCGCTAGGAAAGAAAGTCCGCGTTAAGTGC 82  
 ::::|||||: ||| :|||:|||||:|  
 Db 294 ValGluIleAsnCysThrArgProAsnAsnAsnThrArgLysSerValArgIleGlyPro 313  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 83 AAC-----ACCGATGACACATCGCGGGACTTGAAGAAACTGATAGCGGCC 127  
 :|||:|||||: ||| :|||:|||||:|  
 Db 314 GlyGlnAlaPheTyraThrAsnAspIleIleGlyAsnIleArgGlnAlaHisCysAsn 333  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 128 CAACCTGGCACTCGTTGGATAGATC-----GTTCTTAAAGAGTGGTAC----- 172  
 :|||:|||||: ||| :|||:|||||:|  
 Db 334 IleserThrAspArgTrpAsnLysThrLeuGlnGlnValMetCysLysLeuGlnGluHis 353  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 173 -----ACGATT---TTTAAGGACCATGTATCTCTCGGAGATTATGAATCCAC 217  
 :|||:|||||: ||| :|||:|||||:|  
 Db 354 PheProAsnLysThrIleLysPheLysProHisAlaGly---GlyAspLeuGluIle-Th 372  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 218 GATGGGATGAACCTGAGCTTTATTACCACTAGAGGGGAATTCCTCCACCTTGCCCAACC 277  
 :|||:|||||: ||| :|||:|||||:|  
 Db 372 rMet-----HisSerPhe---AsnCysArgGlyGluPhePheTy- 384  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 278 TTGCTTTCTCTCCCATGGCTCATTTAAACACTGTTGTAGATGCTCATTTTAA----- 329  
 :|||:|||||: ||| :|||:|||||:|  
 Db 385 -----CysAsnThrSerAsnLeuPheAsnSe 393  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 330 -ACAATTACATGAATAAAACTTT 353  
 :|||:|||||: ||| :|||:|||||:|  
 Db 393 rThrTyHisAsnAsnGlyThrTy 401  
 :|||:|||||: ||| :|||:|||||:|

RESULT 2  
 US-11-100-356-3  
 ; Sequence 3, Application US/11100356  
 ; Publication No. US2006057115A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZUR MEDEFE, Jan  
 ; APPLICANT: BARNETT, Susan  
 ; APPLICANT: LIAN, Ying  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B  
 ; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
 ; FILE REFERENCE: 2300-1621.20  
 ; CURRENT APPLICATION NUMBER: US/11/100,356  
 ; CURRENT FILING DATE: 2005-04-06  
 ; PRIOR APPLICATION NUMBER: US/10/190,434  
 ; PRIOR FILING DATE: 2002-07-05  
 ; NUMBER OF SEQ ID NOS: 68  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 867  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: TV1.8\_2  
 US-11-100-356-3

Alignment Scores:  
 Pred. No.: 1.41 Length: 867  
 Score: 72.50 Matches: 40  
 Percent Similarity: 41.3% Conservative: 12  
 Best Local Similarity: 31.7% Mismatches: 32  
 Query Match: 10.9% Indels: 42  
 DB: 7 Gaps: 10

US-10-067-832D-1 (1-370) x US-11-100-356-3 (1-867)

Qy 32 ATTGAGTGGTTTGC-----AACGACCGCTAGGAAAGAAAGTCCGCGTTAAGTGC 82  
 ::::|||||: ||| :|||:|||||:|  
 Db 299 ValGluIleAsnCysThrArgProAsnAsnAsnThrArgLysSerValArgIleGlyPro 318  
 :|||:|||||: ||| :|||:|||||:|  
 Qy 83 AAC-----ACCGATGACACATCGCGGGACTTGAAGAAACTGATAGCGGCC 127  
 :|||:|||||: ||| :|||:|||||:|  
 Db 319 GlyGlnAlaPheTyraThrAsnAspValIleGlyAsnIleArgGlnAlaHisCysAsn 338  
 :|||:|||||: ||| :|||:|||||:|

```
Score: 69.50 Matches: 31
Percent Similarity: 46.5% Conservative: 16
Best Local Similarity: 30.7% Mismatches: 30
Query Match: 10.5% Indels: 24
DB: 5 Gaps: 6

US-10-067-832D-1 (1-370) x US-09-978-360A-460 (1-135)
QY 265 TGGAGGAATTCCTCTACTGTAATAAGCTCCAGGTTTCATCCATCGTGGATTTCATA 206
Db 7 TTPARGAEPYrLYeValLeuValMetVal-----ProLeuValGlyLeuile 23
QY 205 ATCTCCAGAGATACATGCTCTTAATAAATCGTGATACCACTTTT-----AAGAAC 155
Db 24 HisLeuGlyTrpTYrArgileYserSer-----ProValPheGlnLeProLysAsn 41
QY 154 GATCTATTCAACAGAGCCAGTTGGGC-----CGCTATCAGTTCTTCAAGTCCCC 101
Db 42 AspAspIleProGluGlnAspSerLeuGlyLeuSerAsnLeuGlnLysSerGlnIleGln 61
QY 100 GATGGTGTCTCGGTGTT-----GCACTTAACGCGGACTTTCTT 62
Db 62 GlyLys***AlaGlyLeuGlnSerSerGlyLysGluAlaAlaLeuAsnLeuSerPheile 81
QY 61 TCCTAGACGGTCGTTGCCAACCACTCA-----ATCATTTGGCTGG 20
Db 82 Ser-LysGluGluMetLysAsnThrSerTrpIleArgLysAsnTrpLeuLeuValAlaGl 101
QY 19 A 19
Db 101 Y 101

RESULT 4
US-11-100-356-4
; Sequence 4, Application US/11100356
; Publication No. US20060057115A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MESEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/11/100,356
; CURRENT FILING DATE: 2005-04-06
; PRIOR FILING DATE: US/10/190,434
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TV1.8_5
US-11-100-356-4
Alignment Scores:
Pred. No.: 5.69 Length: 869
Score: 67.50 Matches: 33
Percent Similarity: 47.4% Conservative: 13
Best Local Similarity: 34.0% Mismatches: 27
Query Match: 10.2% Indels: 24
DB: 7 Gaps: 8

US-10-067-832D-1 (1-370) x US-11-100-356-4 (1-869)
QY 32 ATTGAGTGGTTTC-----AACACGGCTTAGAAGAAAGTCCGCGTTAAATGTC 82
Db 299 ValGluIleAsnCysThrArgProAsnAsnAsnThrArgLysSerValArgIleGlyPro 318
QY 83 AAC-----ACCGATGACACCATCGGGGACTTGAGAAACTGATAGCGGCC 127
Db 83 AAC-----ACCGATGACACCATCGGGGACTTGAGAAACTGATAGCGGCC 127
```

```
Db 319 GlyGlnAlaPheTYrAlaThrAsnAspValIleGlyAsnIleArgGlnAlaHisCysAsn 338
QY 128 CAAACTGGCACTCGTTGGAATAAGATC-----GTTCTTAAAGTGG----- 169
Db 339 IleSerThrAspArgTIPAsnLysThrLeuGlnGlnValMetLysLysLeuGlyGluHis 358
QY 170 -----TACACGATT-----TTTAAGGACCATGTATCTCTGGGAGATATGAATCCAC 217
Db 359 PheProAsnLysThrIleLysPheGluProHisAlaGly---GlyAspLeuGluile-Th 377
QY 218 GATGGATGAACCTGGAGCTTTATTACCAGTAGAGGGAATTCCTCCAC 266
Db 377 rMet-----HisSerPhe---AsnCysArgGlyGluPhePheTYr 389

RESULT 5
US-11-096-568A-33836
; Sequence 33836, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33836
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(315)
; OTHER INFORMATION: Ceres Seq. ID no. 13604301
US-11-096-568A-33836
Alignment Scores:
Pred. No.: 6.4 Length: 315
Score: 66.50 Matches: 23
Percent Similarity: 47.6% Conservative: 17
Best Local Similarity: 27.4% Mismatches: 27
Query Match: 10.0% Indels: 17
DB: 7 Gaps: 5

US-10-067-832D-1 (1-370) x US-11-096-568A-33836 (1-315)
QY 235 CTCCAGGTTCATCCCATCGTGGATTTCATAATCTCCAGAGATACATCGTCTTAAAAAT 176
Db 16 IleArgValHisProArgMetMetPheTYrAsnValGluLys-----IleGluGluPro 33
QY 175 CGTGATACCACTTTTAAAGAACGATCTTTATTTCCAACGAGTGCACGTTTGGCGCTATCAG 116
Db 34 LysLeuArgPhePheLys---AspIleGlyPheThrGlySerGlyLeuGlyLys----- 50
QY 115 TTCTTCAAGTCCCGGATGGTGTCATCGTGGTTCACCTT-----AAC 74
Db 51 PheValSerGlnAsnSerSerValValGlyValSerLeuValLysLeuIleProThr 70
QY 73 GCGGACTTTCTTCTTAGACGGTGTGTCGAAACCCAC-----CTCAAT 32
Db 71 ValGluLeuLeuLysSerIleValAlaProLysHisGluAspLeuProValIleLeuSer 90
QY 31 CATTTGGCTGG 20
Db 91 ArgCysGlyTrp 94

RESULT 6
US-11-096-568A-33835
; Sequence 33835, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
```

```
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33835
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: Ceres Seq. ID no. 13604300
US-11-096-568A-33835

Alignment Scores:
Pred. No.: 6.45 Length: 332
Score: 66.50 Matches: 23
Percent Similarity: 47.6% Conservative: 17
Best Local Similarity: 27.4% Mismatches: 27
Query Match: 10.0% Indels: 17
DB: 7 Gaps: 5

US-10-067-832D-1 (1-370) x US-11-096-568A-33835 (1-332)
QY 235 CTCGAGTTCATCCCATCGTGGATTTCATAATCTCCAGAGATACATGTCCTTAAAAAT 176
DB 33 IleArgValHisProArgMetMetPheTyAsnValGluLys-----IleLeuGluPro 50
QY 175 CGTGATACCACCTTTTAAAGAACGATCTTATTCCAACGAGTGCCAGTTTGGCCGCTATCAG 116
DB 51 LysLeuArgPhePheLys---AspIleGlyPheThrGlySerGlyLeuGlyLys----- 67
QY 115 TTCTTCAAGTCCCGATGTCATCGGTGTCACCTT-----AAC 74
DB 68 PheValSerGlnAsnSerSerValValGlyValSerLeuValLysLeuLysLeuProThr 87
QY 73 GCGGACTTCTTCTTAGACGGTGGTTCGAAACAC-----CTCAAT 32
DB 88 ValGluLeuLysSerIleValAlaProLysHisGluAspLeuProValIleLeuSer 107
QY 31 CATTTGGCTGG 20
DB 108 ArgCysGlyTrp 111

RESULT 7
US-11-096-568A-33834
; Sequence 33834, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33834
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(391)
; OTHER INFORMATION: Ceres Seq. ID no. 13604299
US-11-096-568A-33834

Alignment Scores:
Pred. No.: 6.62 Length: 391
Score: 66.50 Matches: 23
Percent Similarity: 47.6% Conservative: 17
Best Local Similarity: 27.4% Mismatches: 27
Query Match: 10.0% Indels: 17
```

```
DB: 7 Gaps: 5
US-10-067-832D-1 (1-370) x US-11-096-568A-33834 (1-391)
QY 235 CTCGAGTTCATCCCATCGTGGATTTCATAATCTCCAGAGATACATGTCCTTAAAAAT 176
DB 92 IleArgValHisProArgMetMetPheTyAsnValGluLys-----IleLeuGluPro 109
QY 175 CGTGATACCACCTTTTAAAGAACGATCTTATTCCAACGAGTGCCAGTTTGGCCGCTATCAG 116
DB 110 LysLeuArgPhePheLys---AspIleGlyPheThrGlySerGlyLeuGlyLys----- 126
QY 115 TTCTTCAAGTCCCGATGTCATCGGTGTCACCTT-----AAC 74
DB 127 PheValSerGlnAsnSerSerValValGlyValSerLeuValLysLeuLysLeuProThr 146
QY 73 GCGGACTTCTTCTTAGACGGTGGTTCGAAACAC-----CTCAAT 32
DB 147 ValGluLeuLysSerIleValAlaProLysHisGluAspLeuProValIleLeuSer 166
QY 31 CATTTGGCTGG 20
DB 167 ArgCysGlyTrp 170

RESULT 8
US-11-182-016-31
; Sequence 31, Application US/11182016
; Publication No. US20060019294A1
; GENERAL INFORMATION:
; APPLICANT: SUGEN, INC.
; TITLE OF INVENTION: TYROSINE KINASE SUBSTRATE (TKS) PROTEINS
; FILE REFERENCE: 038602/0102
; CURRENT APPLICATION NUMBER: US/11/182,016
; CURRENT FILING DATE: 2005-07-15
; PRIOR APPLICATION NUMBER: US/09/958,359
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Tks 118
US-11-182-016-31

Alignment Scores:
Pred. No.: 7.91 Length: 495
Score: 66.00 Matches: 30
Percent Similarity: 28.8% Conservative: 6
Best Local Similarity: 24.0% Mismatches: 37
Query Match: 9.9% Indels: 52
DB: 7 Gaps: 7

US-10-067-832D-1 (1-370) x US-11-182-016-31 (1-495)
QY 299 GAGCCATGGGAGAGAGAAAGCAAGTTGGCGAAGTGGAGGAATTCCTCTACTGGTAAT 240
DB 361 LysProTrpPro-----ProSerProCysThr 369
QY 239 AAAGCTCCAGTTCA-----TCCCAT----- 219
DB 370 GlnAlaProHisProProArgProValArgTrpSerHisGlyProProSerGlySerTrp 389
QY 218 -----CGTGGATTTCATAATCTCCAGAGATACAT----- 189
DB 390 ProTrpCysArgGlyTrpHisArgLeuProSerAlaHisArgSerArgProArgLeuSer 409
QY 188 -----GGTCCCTTAAAAATCGGTACCACCTTTTAA 159
DB 410 SerGlyGlnIleTrpAlaValGlnSerTrpGlyPro---SerLeuCysArg----- 425
QY 158 GAACGATCTTATTCCAACGAGTGCCAGTTTGGCGCGCTATCAGTTTCT----- 111
```

```
Db 426 ArgArgThrSerProSerArgCysAlaProProProSerProProGlyHisProProLeu 445
Qy 110 TCAAGTCCCGATGGTGCATCGGTGGCTAAACCGGACTTCTTCTAGACGGT 51
Db 446 CysGlnProArgGlyCysHisCysCysLeuHisArgGluProSerArgSerGly 465
Qy 50 CGTTGCAAAACACCT 36
Db 466 ThrSerArgProPro 470
RESULT 9
US-10-793-626-3104
; Sequence 3104, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3104
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3104
Alignment Scores:
Pred. No.: 8.45 Length: 312
Score: 65.50 Matches: 26
Percent Similarity: 37.5% Conservative: 22
Best Local Similarity: 20.3% Mismatches: 33
Query Match: 9.8% Indels: 47
DB: 6 Gaps: 5
US-10-067-832D-1 (1-370) x US-10-793-626-3104 (1-312)
Qy 334 ATTGTTAAATAAT-----GAGCATCTACACAGAGTGTAAATGAGCCAT----- 293
Db 153 IleVallysGlyTyrPheAlaAspGlnLeuGlnGlnTyrPheLeuAlaHisGlyValSer 172
Qy 292 -----GGGAGAGGNAAGCAAGGTGG 272
Db 173 SerGlylleIleAspLeuGlyGlyAsnValLeuThrIleGlyArgGlnProGluThrLeu 192
Qy 271 GCAAGGTGG-----AGGAATTCC-----CCTCTACTGGTA 242
Db 193 GluIysTrpHisValGlyValArgAsnProPheHisIysAspThrLeuProLeuValThr 212
Qy 241 ATAAAGTCCAGGTTCATCCCATCGTGATTTCAATATCTCCAGAGATACGGTCTT 182
Db 213 LeuSerValGluHisGlnSerValValThrSerGlyIleTyrGluArgTyrPheIleGln 232
Qy 181 AAAAATCGTGTACCACCTTTTAAAGACGATCTTATCCACGAGTGCAGTTGGCGGC 122
Db 233 GluAsnGlnLeu----- 236
Qy 121 TATCAGTGTCTTCAAGTCCCGATGGTGTATCGGTGTGCATCTTAACGGGACTTTCTT 62
Db 237 -----PheHisIleLeuAspSerThrThrGlyTyrProValAspAsnAspIleAla 254
Qy 61 TCCTAGACGGTGGTTGCAACACCAC 38
Db 255 SerValThrIleIleSerAspHis 262
RESULT 10
```

```
US-11-087-099-6646
; Sequence 6646, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6646
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-11-087-099-6646
Alignment Scores:
Pred. No.: 8.69 Length: 371
Score: 65.50 Matches: 28
Percent Similarity: 49.4% Conservative: 16
Best Local Similarity: 31.5% Mismatches: 35
Query Match: 9.8% Indels: 10
DB: 7 Gaps: 4
US-10-067-832D-1 (1-370) x US-11-087-099-6646 (1-371)
Qy 262 AGGAATCCCTCTACTGTGTAATAAAGCTCCAGGTTTCATCCCATCGTGGATTTTCAATAATC 203
Db 205 ArgAsnGluProLeuLeuProValPheLeuGluTrpThrProIleValValPheThrLeu 224
Qy 202 TCCAGAGATACATGGTCCCTTAAAAATCGGTACACCTTTTAAAGACGATCTTATTCCCA 143
Db 225 SerVal-GlyThrTrp-----LeuTyrSerProTyrSerThrLeuMetAspG1 240
Qy 142 ACGA-----GTGCCAGTTTGGCGCGTATCAGTTTCTTCAAGTCCCGATGGTGTATC 89
Db 240 uAsnHisLeuValLeuPheCysValThrMetSerPheValPheGlyArgMetThrTrpLy 260
Qy 88 GGTGTG-----CACTTAACGGCGACTTTTCTTAGACGGTCTGTGGCAACACCCTC 35
Db 260 sMetIleLeuAlaHisLeuThrLysGlnProPheProTyrTrpThrValMetLeuTrpPr 280
Qy 34 AATCATTTGCTGGCTGACCTGTAATC 10
Db 280 oLeuile---GlyGlyAlaLeuile 287
RESULT 11
US-11-119-249-7
; Sequence 7, Application US/11119249
; Publication No. US20050261190A1
; GENERAL INFORMATION:
; APPLICANT: SK Corp.
; APPLICANT: Ewha Womans University
; APPLICANT: LEE, Kong-Joo
; APPLICANT: KIM, Hee-Jung
; APPLICANT: CHO, Jeong Woo
; APPLICANT: KIM, Eunhee
; APPLICANT: SONG, Eun Joo
; APPLICANT: MAENG, Cheol Young
; TITLE OF INVENTION: FAS ASSOCIATED FACTOR 1
; FILE REFERENCE: 12120-02USA
; CURRENT APPLICATION NUMBER: US/11/119,249
; CURRENT FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: US 60/566,966
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: US 60/590,327
; PRIOR FILING DATE: 2004-07-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 7
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
```





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 23:55:38 ; Search time 190 Seconds  
(without alignments)  
168.814 Million cell updates/sec

Title: US-10-067-832d-2  
Perfect score: 394  
Sequence: 1 MIEVVCNDRLGKVRVKNT.....HVSIGDYEIHDGMNLELYQ 73

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	100.0	73	2 AAY08413	Aay08413 P. obesu
2	394	100.0	73	3 AAB36290	Aab36290 Israeli s
3	394	100.0	73	4 AAM39680	Aam39680 Human pol
4	394	100.0	89	4 AAM41466	Aam41466 Human pol
5	394	100.0	115	5 ABP41331	Abp41331 Human ova
6	339	86.0	73	4 ABB59910	Abb59910 Drosophil
7	313	79.4	73	3 AAG41925	Aag41925 Arabidops
8	313	79.4	96	3 AAG41924	Aag41924 Arabidops
9	312	79.2	73	3 AAG40885	Aag40885 Zea mays
10	300	76.1	73	3 AAG27281	Aag27281 Zea mays
11	285	72.3	73	8 ADT58384	Adt58384 Plant pol
12	252	64.0	89	5 ADH23212	Adh23212 Novel yea
13	252	64.0	320	6 ABR83620	Abr83620 HUB1-GFP
14	163	41.4	33	2 AAY08414	Aay08414 Human bea
15	163	41.4	33	8 ADR43346	Adr43346 Human bea
16	147.5	37.4	32	3 AAB36291	Aab36291 Human bea
17	138	35.0	27	4 AAM21187	Aam21187 Peptide #
18	138	35.0	27	4 ABB34508	Abb34508 Peptide #
19	138	35.0	27	4 AAM37402	Aam37402 Peptide #
20	138	35.0	27	4 ABB26468	Abb26468 Protein #
21	138	35.0	27	4 AAM77251	Aam77251 Human bon
22	138	35.0	27	4 AAM64443	Aam64443 Human bra
23	138	35.0	27	5 ABG46267	Abg46267 Human pep
24	137	34.8	40	9 ADV95411	Adv95411

25	133	33.8	40	9 ADV95410	Adv95410 Yeast ubi
26	120	30.5	32	9 ADV95409	Adv95409 Yeast ubi
27	115	29.2	32	9 ADV95412	Adv95412 Yeast ubi
28	85.5	21.7	857	9 ADX39713	Adx39713 HIV Env p
29	81	20.6	638	9 ADZ07768	Adz07768 HIV CON-A
30	81	20.6	854	9 ADZ07813	Adz07813 HIV CON-0
31	81	20.6	854	9 ADZ07767	Adz07767 HIV CON-A
32	80.5	20.4	857	9 ADX39712	Adx39712 HIV Env p
33	80	20.3	477	3 AAG38562	Aag38562 Arabidops
34	80	20.3	553	3 AAG38561	Aag38561 Arabidops
35	80	20.3	631	3 AAG38560	Aag38560 Arabidops
36	80	20.3	853	9 ADZ04172	Adz04172 Env prote
37	80	20.3	854	9 ADX39709	Adx39709 HIV Env p
38	76	19.3	493	9 ADM69641	Adm69641 HIV-1 gp1
39	76	19.3	831	9 ADZ07806	Adz07806 HIV CON-F
40	75.5	19.2	876	9 ADX39742	Adx39742 HIV Env p
41	74.5	18.9	509	9 ADM69547	Adm69547 HIV-1 gp1
42	74.5	18.9	630	9 ADZ07780	Adz07780 HIV subty
43	74.5	18.9	854	9 ADZ07779	Adz07779 HIV Env p
44	74.5	18.9	857	9 ADX39691	Adx39691 HIV Env p
45	74.5	18.9	1066	8 ADM92082	Adm92082 S pneumon

## ALIGNMENTS

## RESULT 1

AAY08413  
ID AAY08413 standard; protein; 73 AA.  
XX  
AC AAY08413;  
XX  
DT 24-JUL-1999 (first entry)  
XX  
DE P. obesu beacon protein.  
XX  
KW Beacon, hypothalamus; obese; lean; agonist; antagonist; treatment;  
KW obesity; anorexia; weight maintenance; energy imbalance; diabetes;  
KW metabolic syndrome; dyslipidemia; hypertension; insulin resistance;  
KW medicament; livestock; diagnosis.  
XX  
OS Psammomys obesus.  
XX  
PN WO9923217-A1.  
XX  
PD 14-MAY-1999.  
XX  
PP 30-OCT-1998; 98WO-AU000902.  
XX  
PR 31-OCT-1997; 97AU-00000117.  
PR 11-NOV-1997; 97AU-00000323.  
PA (ITDI-) INT DIABETES INST.  
PA (UYDE-) UNIV DEAKIN.  
XX  
PI Zimmet PZ, Collier G;  
DR WPI; 1999-337484/28.  
DR N-PSDB; AAX57359.  
XX  
PT New gene encoding a beacon protein associated with modulation of obesity,  
PT diabetes and metabolic energy levels.  
XX  
PS Claim 2; Page 50; 85pp; English.  
XX  
CC This invention describes a novel beacon protein and its encoding nucleic  
CC acid which is expressed in larger amounts in hypothalamus tissue of obese  
CC animals compared to lean animals. Agonists and antagonists of beacon can  
CC be used to treat obesity, anorexia, weight maintenance, energy imbalance,  
CC diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin  
CC resistance. The beacon protein, itself is used to manufacture medicaments  
CC for treatment of obesity, anorexia, energy imbalance or diabetes. The  
CC treatment is contemplated for both human and animals, such as those

CC important to the livestock industry. The antibody and polynucleotides are  
CC useful in diagnosis of conditions as above  
XX  
SQ Sequence 73 AA;

Query Match 100.0%; Score 394; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 7.5e-42;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60  
DB 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYQ 73  
DB 61 EIHGGMNLELYQ 73

RESULT 2  
AAB36290  
ID AAB36290 standard; protein; 73 AA.

XX AAB36290;  
XX AC  
XX XX  
XX 23-FEB-2001 (first entry)  
XX DE Israeli sand rat beacon ligand.  
XX XX  
XX Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;  
XX KW ligand.

XX OS Psammomys obesus.  
XX XX  
XX WO200064931-A1.  
XX PN  
XX 02-NOV-2000.

XX PF 19-APR-2000; 2000WO-AU000342.  
XX XX  
XX 23-APR-1999; 99AU-00009919.  
XX PR 24-MAR-2000; 2000AU-00006454.

XX XX  
XX (AUTO-) AUTOGEN PTY LTD.  
XX PA  
XX Collier G, Walder K, Zimmet P;  
XX PI  
XX WPI; 2000-687311/67.  
XX DR N-PSDB; AAC81767.

XX XX  
XX Ligand of beacon protein useful for treating obesity, anorexia, energy  
XX PT imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and  
XX PT insulin resistance.

XX PS Claim 3; Fig 1; 67pp; English.

XX XX  
XX The present invention is related to the isolation of a ligand known as  
XX CC beacon from the Israeli sand rat. Beacon is associated with the  
XX CC regulation of energy balance, and the protein, its coding sequence and  
XX CC analogues can be used in the treatment of diabetes, obesity, anorexia,  
XX CC energy imbalance, metabolic syndrome, dyslipidaemia, hypertension and  
XX CC insulin resistance. In addition, they can be used in agriculture to  
XX CC produce leaner animals

XX SQ Sequence 73 AA;

Query Match 100.0%; Score 394; DB 3; Length 73;  
Best Local Similarity 100.0%; Pred. No. 7.5e-42;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60  
DB 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYQ 73  
DB 61 EIHGGMNLELYQ 73

RESULT 3  
AAM39680  
ID AAM39680 standard; protein; 73 AA.

XX AAM39680;  
XX AC  
XX 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2825.

XX XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;  
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-0052317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX PI Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR N-PSDB; AAI58836.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such  
XX PT as central nervous system injuries.

XX PS Example 4; SEQ ID NO 2825; 10078pp; English.

XX XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide  
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and  
XX CC localised neuropathies and central nervous system diseases, such as  
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX CC utilisation of the activities such as: Immune system suppression,  
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and  
XX CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX CC part of the printed specification

XX SQ Sequence 73 AA;

Query Match 100.0%; Score 394; DB 4; Length 73;

CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 89 AA;

Query Match 100.0%; Score 394; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 9.6e-42;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVVCNDRGLGKVKVCKNTDDTIGDLKKLIAAQTGTRNNKIVLKKWYTIKDHVSLGDY 60  
Db 17 MIEVVCNDRGLGKVKVCKNTDDTIGDLKKLIAAQTGTRNNKIVLKKWYTIKDHVSLGDY 76

QY 61 EIHDMNLELYQ 73  
Db 77 EIHDMNLELYQ 89

RESULT 5  
ABP41331  
ID ABP41331 standard; protein; 115 AA.  
AC ABP41331;  
XX  
XX 22-AUG-2002 (first entry)  
XX Human ovarian antigen HACNC39, SEQ ID NO:2463.  
DE  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
OS Homo sapiens.  
XX  
XX WO20020677-A1.  
XX  
XX 03-JAN-2002.  
XX  
XX 07-JUN-2001; 2001WO-US018569.  
XX  
XX 07-JUN-2000; 2000US-0209467P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
XX  
XX WPI; 2002-147878/19.  
XX  
XX N-PSDB; ABQ54408.  
XX  
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.  
XX  
XX Claim 11; SEQ ID NO 2463; 2922pp; English.  
XX  
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC

metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA
CC	sequences (ABU01840-ABL16175) and the encoded proteins (ABB57737-
CC	ABB72072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at fip.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 73 AA;
	Query Match                86.0%; Score 339; DB 4; Length 73;
	Best Local Similarity      86.3%; Pred.No. 6.9e-35;
	Matches 63; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy	1 MIEVVCNDRLGKKVRKCNFTDITGLKKLIAAQTGTNRNKIVLKKWYTFKDHSVLGDY 60
Dd	1 MIEITTCNDRLGKKVRKCNPDITGLKKLIAAQTGTKEKIVLKWKYTIFKDIPLSDY 60
Qy	61 EIHGDNWLELYQ 73
Dd	61 EIHGDNWLELYQ 73
RESULT 7	
AAG41925	
ID	AAG41925 standard; protein; 73 AA.
XX	
AC	AAG41925;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 52222.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	23-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135639P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.

```
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 79.4%; Score 313; DB 3; Length 73;
Best Local Similarity 80.6%; Pred. No. 1.4e-31;
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGRWNKIVLKKWYTFKDHVSLG DY 60
DB 1 MIEVVLNDRLGKVRKVCNDDDTIGDLKKLVAAQTGTAEKIRIQKWNYYKDHITL KDY 60

QY 61 EIHGGMNLELY 72
DB 61 EIHGGMGLELY 72

RESULT 8
AAG41924
ID AAG41924 standard; protein; 96 AA.
XX
AC AAG41924;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52221.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
```

```
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148694P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159631P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161599P.

PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 79.4%; Score 313; DB 3; Length 96;
Best Local Similarity 80.6%; Pred. No. 1.9e-31;
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRIGKKVRKVCNTDDTIGDLKKLIAAQTGRWKKIVLKKWYTFPKDHVSLG DY 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 MIEVLCNDRIGKKVRKVCNTDDTIGDLKKLIAAQTGRWKKIVLKKWYTFPKDHVSLG DY 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 EIHGGMNLELYY 72
    ||||| ||||| |||||
Db 84 EIHGGMNLELYY 95
    ||||| ||||| |||||

RESULT 9
AAG40885
ID AAG40885 standard; protein; 73 AA.
XX
AC AAG40885;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 50790.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
25-FEB-2000; 2000EP-00301439.
XX
25-FEB-1999; 99US-0121825P.
XX
03-MAR-1999; 99US-0123180P.
XX
03-MAR-1999; 99US-0123548P.
XX
23-MAR-1999; 99US-0125788P.
XX
25-MAR-1999; 99US-0126264P.
XX
25-MAR-1999; 99US-0126785P.
XX
01-APR-1999; 99US-0127462P.
XX
06-APR-1999; 99US-0128234P.
XX
08-APR-1999; 99US-0128714P.
XX
16-APR-1999; 99US-0129845P.
XX
19-APR-1999; 99US-0130077P.
XX
21-APR-1999; 99US-0130449P.
XX
23-APR-1999; 99US-0130510P.
XX
23-APR-1999; 99US-0130891P.
XX
28-APR-1999; 99US-0131449P.
XX
30-APR-1999; 99US-0132048P.
XX
30-APR-1999; 99US-0132407P.
XX
04-MAY-1999; 99US-0132484P.
XX
05-MAY-1999; 99US-0132485P.
XX
06-MAY-1999; 99US-0132486P.
XX
06-MAY-1999; 99US-0132487P.
XX
07-MAY-1999; 99US-0132863P.
XX
11-MAY-1999; 99US-0134256P.
XX
14-MAY-1999; 99US-0134218P.
XX
14-MAY-1999; 99US-0134219P.
XX
14-MAY-1999; 99US-0134221P.
XX
18-MAY-1999; 99US-0134370P.
XX
18-MAY-1999; 99US-0134768P.
XX
19-MAY-1999; 99US-0134941P.
XX
20-MAY-1999; 99US-0135124P.
XX
21-MAY-1999; 99US-0135353P.
```

PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 26-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 02-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 05-JUN-1999; 99US-0138094P.  
PR 06-JUN-1999; 99US-0138540P.  
PR 07-JUN-1999; 99US-0138847P.  
PR 08-JUN-1999; 99US-0139119P.  
PR 09-JUN-1999; 99US-0139452P.  
PR 10-JUN-1999; 99US-0139453P.  
PR 11-JUN-1999; 99US-0139492P.  
PR 12-JUN-1999; 99US-0139454P.  
PR 13-JUN-1999; 99US-0139455P.  
PR 14-JUN-1999; 99US-0139456P.  
PR 15-JUN-1999; 99US-0139457P.  
PR 16-JUN-1999; 99US-0139458P.  
PR 17-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 19-JUN-1999; 99US-0139461P.  
PR 20-JUN-1999; 99US-0139462P.  
PR 21-JUN-1999; 99US-0139463P.  
PR 22-JUN-1999; 99US-0139464P.  
PR 23-JUN-1999; 99US-0139465P.  
PR 24-JUN-1999; 99US-0139466P.  
PR 25-JUN-1999; 99US-0139467P.  
PR 26-JUN-1999; 99US-0139468P.  
PR 27-JUN-1999; 99US-0139469P.  
PR 28-JUN-1999; 99US-0139470P.  
PR 29-JUN-1999; 99US-0139471P.  
PR 30-JUN-1999; 99US-0140287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 03-JUL-1999; 99US-0142055P.  
PR 04-JUL-1999; 99US-0142390P.  
PR 05-JUL-1999; 99US-0142803P.  
PR 06-JUL-1999; 99US-0143220P.  
PR 07-JUL-1999; 99US-0143277P.  
PR 08-JUL-1999; 99US-0143542P.  
PR 09-JUL-1999; 99US-0143624P.  
PR 10-JUL-1999; 99US-0144005P.  
PR 11-JUL-1999; 99US-0144085P.  
PR 12-JUL-1999; 99US-0144086P.  
PR 13-JUL-1999; 99US-0144325P.  
PR 14-JUL-1999; 99US-0144331P.  
PR 15-JUL-1999; 99US-0144332P.  
PR 16-JUL-1999; 99US-0144333P.  
PR 17-JUL-1999; 99US-0144334P.  
PR 18-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144684P.  
PR 22-JUL-1999; 99US-0144814P.  
PR 23-JUL-1999; 99US-0145086P.  
PR 24-JUL-1999; 99US-0145088P.  
PR 25-JUL-1999; 99US-0145085P.  
PR 26-JUL-1999; 99US-0145087P.  
PR 27-JUL-1999; 99US-0145089P.  
PR 28-JUL-1999; 99US-0145192P.  
PR 29-JUL-1999; 99US-0145145P.  
PR 30-JUL-1999; 99US-0145218P.  
PR 31-JUL-1999; 99US-0145224P.  
PR 01-AUG-1999; 99US-0145276P.  
PR 02-AUG-1999; 99US-0145913P.  
PR 03-AUG-1999; 99US-0145916P.  
PR 04-AUG-1999; 99US-0145918P.  
PR 05-AUG-1999; 99US-0145919P.  
PR 06-AUG-1999; 99US-0145951P.  
PR 07-AUG-1999; 99US-0146386P.  
PR 08-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 06-AUG-1999; 99US-0147192P.  
PR 07-AUG-1999; 99US-0147260P.  
PR 08-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 10-AUG-1999; 99US-0147493P.  
PR 11-AUG-1999; 99US-0147935P.  
PR 12-AUG-1999; 99US-0148171P.  
PR 13-AUG-1999; 99US-0148319P.  
PR 14-AUG-1999; 99US-0148341P.  
PR 15-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 19-AUG-1999; 99US-0149262P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 21-AUG-1999; 99US-0149722P.  
PR 22-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 24-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 28-AUG-1999; 99US-0151080P.  
PR 29-AUG-1999; 99US-0151303P.  
PR 30-AUG-1999; 99US-0151338P.  
PR 31-AUG-1999; 99US-0151930P.  
PR 01-SEP-1999; 99US-0152363P.  
PR 02-SEP-1999; 99US-0153070P.  
PR 03-SEP-1999; 99US-0153758P.  
PR 04-SEP-1999; 99US-0154018P.  
PR 05-SEP-1999; 99US-0154039P.  
PR 06-SEP-1999; 99US-0154779P.  
PR 07-SEP-1999; 99US-0155139P.  
PR 08-SEP-1999; 99US-0155486P.  
PR 09-SEP-1999; 99US-0155659P.  
PR 10-SEP-1999; 99US-0156458P.  
PR 11-SEP-1999; 99US-0156596P.  
PR 12-SEP-1999; 99US-0157117P.  
PR 13-SEP-1999; 99US-0157753P.  
PR 14-SEP-1999; 99US-0157865P.  
PR 15-SEP-1999; 99US-0158029P.  
PR 16-SEP-1999; 99US-0158232P.  
PR 17-SEP-1999; 99US-0158369P.  
PR 18-SEP-1999; 99US-0159293P.  
PR 19-SEP-1999; 99US-0159294P.  
PR 20-SEP-1999; 99US-0159295P.  
PR 21-SEP-1999; 99US-0159329P.  
PR 22-SEP-1999; 99US-0159330P.  
PR 23-SEP-1999; 99US-0159331P.  
PR 24-SEP-1999; 99US-0159637P.  
PR 25-SEP-1999; 99US-0159638P.  
PR 26-SEP-1999; 99US-0159584P.  
PR 27-SEP-1999; 99US-0160741P.  
PR 28-SEP-1999; 99US-0160767P.  
PR 29-SEP-1999; 99US-0160768P.  
PR 30-SEP-1999; 99US-0160770P.  
PR 01-OCT-1999; 99US-0160814P.  
PR 02-OCT-1999; 99US-0160815P.  
PR 03-OCT-1999; 99US-0160980P.  
PR 04-OCT-1999; 99US-0160981P.  
PR 05-OCT-1999; 99US-0160989P.  
PR 06-OCT-1999; 99US-0161404P.  
PR 07-OCT-1999; 99US-0161405P.  
PR 08-OCT-1999; 99US-0161406P.  
PR 09-OCT-1999; 99US-0161359P.  
PR 10-OCT-1999; 99US-0161360P.  
PR 11-OCT-1999; 99US-0161361P.  
PR 12-OCT-1999; 99US-0161920P.



```
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match          79.2%; Score 312; DB 3; Length 73;
Best Local Similarity 80.6%; Pred. No. 1.8e-31;
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRLGKKVRVKNTDDTTGDLKKLIAAQTGRNKKIVLKKWYTFKDHVSLGDY 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MIEVVLNDRLGKKVRVKCNEDDTGDLKKLVAAGTGRPEKIRIQKWNVIYKDHITLKDY 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 EIHGGMNLELY 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 EIHGGMNLELY 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AAG27281
ID AAG27281 standard; protein; 73 AA.
XX
AC AAG27281;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 32055.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138034P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
```



CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20040216190.

CC Sequence 73 AA;

Query Match 72.3%; Score 285; DB 8; Length 73;  
Best Local Similarity 76.7%; Pred. No. 4.8e-28;  
Matches 56; Conservative 7; Mismatches 8; Indels 2; Gaps 2;

QY 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKGLIAAQTG-TRWNKIVLKKWYTFKDHVSLGD 59  
DB 1 MIEVVNDRLGKVRVKC-XDDTTIGDLKGLVAAQTGTRPEKIRIQKWNYYIKDHITLKD 59  
QY 60 YEIHGDMNLELY 72  
DB 60 YEVDHGMLELY 72

RESULT 12

ID ADH32312  
AC ADH32312 standard; protein; 89 AA.

XX ADH32312;

DT 11-MAR-2004 (first entry)

XX Novel yeast smORF572-encoded polypeptide, SEQ ID NO:770.

XX Open reading frame; ORF; identification; in silico; yeast; smORF;  
XX small open reading frame; antisense therapy; antibody therapy;  
XX drug screening; fungal infection; fungicide; gene therapy; vaccine.

XX Saccharomyces cerevisiae.

OS WO200268693-A2.

PN 06-SEP-2002.

PD 27-FEB-2002; 2002WO-US005677.

PF 27-FEB-2001; 2001US-0271406P.

PR 29-NOV-2001; 2001US-0333726P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Zeng Q, Kessler M, Cottarel G;

XX WPI; 2002-707018/76.

DR N-PSDB; ADH31639.

XX Identifying open reading frames (ORFs) in a genome of an organism by-  
XX comparing genomic sequences to one or more genomic libraries containing  
XX ORFs, useful for diagnosing, preventing and/or treating fungal  
XX infections.

XX Claim 38; SEQ ID NO 770; 160pp; English.

XX The invention relates to a method of identifying open reading frames

CC (ORFs) in a genome of an organism. The method comprises: collecting a  
CC genomic sequence of a first organism; comparing the genomic sequence of  
CC the first organism to one or more other genomic libraries comprising  
CC genomes of other organisms containing ORFs; and determining ORFs for the  
CC first organism based on the comparison. The invention also relates to  
CC smORFs (small open reading frames) from the Saccharomyces cerevisiae  
CC genome which encode polypeptides less than 100 amino acids long; vectors  
CC and host cells comprising a smORF; polypeptides encoded by the smORFs;  
CC antisense compounds targeted to the smORFs and methods for their use in  
CC inhibiting smORF expression; a method of identifying an inhibitor of a  
CC smORF-encoded protein; an antibody against a smORF-encoded polypeptide;  
CC and a composition comprising a smORF or smORF-encoded polypeptide and a  
CC carrier. The invention provides a useful in silico method of identifying  
CC new coding sequences, including homologues of coding sequences, in yeast  
CC and other organisms. The compositions of the invention, and methods of  
CC their use or discovery are also useful for diagnosing, preventing and/or  
CC treating fungal infections. Sequences ADH32216-ADH32888 represent  
CC specifically claimed yeast smORF-encoded polypeptides. Note: The sequence  
CC data for this patent is not represented in the printed specification but  
CC is based on sequence information supplied to Derwent by the European  
CC Patent Office.

SQ Sequence 89 AA;

Query Match 64.0%; Score 252; DB 5; Length 89;  
Best Local Similarity 65.3%; Pred. No. 9.3e-24;  
Matches 47; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKGLIAAQTGTRWNKIVLKKWYTFKDHVSLGDY 60  
DB 17 MIEVVNDRLGKVRVKCLAE DSVGDFKVLISLIQIGTQPNKIVLQKGGSVLKHISLEDY 76

QY 61 EIHGDMNLELY 72

DB 77 EVHDQTNLELY 88

RESULT 13

ABR83620  
ID ABR83620 standard; protein; 320 AA.

XX ABR83620;

XX 16-OCT-2003 (first entry)

DE HUB1-GFP protein sequence SEQ ID NO:11.

XX Protein expression; purification; fusion protein; SUMO; GFP; RUB; HUB;  
XX green fluorescent protein; enhancing protein expression level; APG8;  
XX AFG12; URM1; ISG15; secretion.

OS Homo sapiens.

OS Asquorea victoria.

OS Synthetic.

XX WO2003057174-A2.

XX 17-JUL-2003.

XX 07-JAN-2003; 2003WO-US000436.

XX 07-JAN-2002; 2002US-0346449P.

XX (LIFE-) LIFESENSORS INC.

XX Butt TR, Weeks SD, Tran HT, Malakhov MP, Malakhova OA;

XX WPI; 2003-577500/54.

DR N-PSDB; ACF57380.

XX Enhancing protein expression levels in or from a cell comprises operably  
XX linking ubiquitin or ubiquitin-like proteins to a nucleic acid encoding a  
XX protein to generate a construct encoding a fusion protein for expression

PT in a host cell.

XX Example 1; Fig 28A-B; 144pp; English.

XX

CC The present invention describes a method for enhancing expression levels

CC of a protein in or from a cell. The method comprises: (a) operably

CC linking a nucleic acid sequence encoding SUMO, RUB, HUB, APG8, APG12,

CC URM1 or ISG15, to a nucleic acid sequence encoding the protein to a

CC generate a construct encoding a fusion protein; and (b) introducing the

CC nucleic acid into a host cell, where the presence of the molecule in the

CC fusion protein increases the expression level or the secretion of the

CC protein in the host cell. The method is useful for enhancing protein

CC expression, and for generating novel amino termini on proteins of

CC interest for a variety of research, diagnostic and therapeutic

CC applications. The method of enhancing protein circumvents problems

CC encountered with previous methods by enhancing the expression of under-

CC expressed proteins, increasing the solubility of proteins that are

CC insoluble, protecting the candidate proteins from degradation by

CC intracellular proteases by fusing ubiquitin-like proteins (UBLs) to their

CC N-termini, cleaving the fusion protein to efficiently generate authentic

CC proteins using naturally present enzymes, generating proteins with novel

CC amino termini, and cleaving all fusion proteins with remarkable

CC efficiency irrespective of the N-terminal sequence of the fused protein,

CC using UBL hydrolases. The present sequence represents a HUB1-GFP protein

CC sequence, which is used in the exemplification of the present invention

XX

SQ Sequence 320 AA;

Query Match 64.0%; Score 252; DB 6; Length 320;

Best Local Similarity 65.3%; Pred. No. 4.7e-23;

Matches 47; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKLLIAAQTGRWNKIVLKKWYIFKDHVSLGDY 60

DB 10 MIEVNVNDRLGKVRKVCNTDDTIGDLKLLIAAQTGRWNKIVLKKWYIFKDHVSLGDY 69

QY 61 EIHGDMNLELY 72

DB 70 EVHDQTNLELY 81

RESULT 14

ID AAY08414

XX AAY08414 standard; protein; 33 AA.

AC AAY08414;

XX

XX 24-JUL-1999 (first entry)

XX Human beacon protein.

DE

XX Beacon; hypothalamus; obese; lean; agonist; antagonist; treatment;

KW obesity; anorexia; weight maintenance; energy imbalance; diabetes;

KW metabolic syndrome; dyslipidemia; hypertension; insulin resistance;

KW medicament; livestock; diagnosis; human.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Misc-difference 15

FT /label= unknown

FT /note= "encoded by CNC"

XX

PN WO9923217-A1.

XX

XX 14-MAY-1999.

XX

XX 30-OCT-1998; 98WO-AU000902.

XX

XX 31-OCT-1997; 97AU-00000117.

PR 11-NOV-1997; 97AU-00000323.

XX

XX (ITDI-) INT DIABETES INST.

PA

(UYDE-) UNIV DEAKIN.

PA Zimmet PZ, Collier G;

XX WPI; 1999-337484/28.

DR N-PSDB; AAX57370.

XX

XX New gene encoding a beacon protein associated with modulation of obesity,

PT diabetes and metabolic energy levels.

XX Claim 2; Page 54; 85pp; English.

XX This invention describes a novel beacon protein and its encoding nucleic

CC acid which is expressed in larger amounts in hypothalamus tissue of obese

CC animals compared to lean animals. Agonists and antagonists of beacon can

CC be used to treat obesity, anorexia, weight maintenance, energy imbalance,

CC diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin

CC resistance. The beacon protein, itself is used to manufacture medicaments

CC for treatment of obesity, anorexia, energy imbalance or diabetes. The

CC treatment is contemplated for both human and animals, such as those

CC important to the livestock industry. The antibody and polynucleotides are

CC useful in diagnosis of conditions as above

XX

SQ Sequence 33 AA;

Query Match 41.4%; Score 163; DB 2; Length 33;

Best Local Similarity 97.0%; Pred. No. 4.9e-13;

Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKLLIAA 33

DB 1 MIEVNCNDRLGKVRKVCNTDDTIGDLKLLIAA 33

RESULT 15

ADR43346

ID ADR43346 standard; peptide; 33 AA.

XX

AC ADR43346;

XX

XX 04-NOV-2004 (first entry)

XX Human Beacon protein.

XX

XX Beacon; Anorectic; Anabolic; Antidiabetic; Cytostatic; Cardiant;

KW Antiinflammatory; Immunosuppressive; Antinfertility; Neuroprotective;

KW Nootropic; CLK; myopathy; obesity; anorexia; diabetes; cancer;

KW heart disease; inflammation; Alzheimer's; Parkinson's; Pearson Syndrome;

KW Pyruvate Carboxylase Deficiency; Pyruvate Dehydrogenase Deficiency;

KW Cancer; Acoustic Neuroma; Acute Lymphocytic Leukemia; Carcinoma;

KW Leukemia.

XX

OS Homo sapiens.

XX

XX WO2004069866-A1.

PN

XX 19-AUG-2004.

XX

XX 10-FEB-2004; 2004WO-AU000147.

PF

XX 10-FEB-2003; 2003US-0446191P.

PR

XX (AUTO-) AUTOGEN RES PTY LTD.

PA (UYDE-) UNIV DEAKIN.

XX

XX Collier G, Walder K, Kerr-Bayles L;

PI

XX WPI; 2004-604412/58.

DR

XX New isolated ligands of mammalian or avian Beacon, useful for e.g.

PT preventing or treating disorders associated with myopathy, obesity,

PT diabetes, cancer, heart disease, inflammation, or disorders associated

PT with the immune system.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 23:59:48 ; Search time 39 Seconds  
(without alignments)  
180.098 Million cell updates/sec

Title: US-10-067-832d-2  
Perfect score: 394  
Sequence: 1 MIEVVCNDRGLGKVRKVCNT.....HVS LGDY EIH DGMNLELYYQ 73

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: Pirl.\*  
2: Pirl.\*  
3: Pirl.\*  
4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	331	84.0	73	2 T25763	hypothetical prote
2	306	77.7	73	2 T47458	hypothetical prote
3	301	76.4	73	2 T40200	ubiquitin-like pro
4	252	64.0	73	2 S78735	protein YNR032c-a
5	80	20.3	631	2 S55243	upiquitin-like pro
6	74.5	18.9	1066	2 B95037	hyaluronidase (imp
7	73.5	18.7	1078	2 F97907	hyaluronate lyase
8	70.5	17.9	299	2 S60529	envelope polyprote
9	70.5	17.9	299	2 S60528	envelope polyprote
10	69	17.5	226	2 T29404	hypothetical prote
11	68.5	17.4	831	2 A70363	mannose-1-phosphat
12	68	17.3	294	2 S60545	envelope polyprote
13	68	17.3	294	2 S60524	envelope polyprote
14	68	17.3	709	2 G96610	probable disease r
15	66.5	16.9	185	2 C97816	aminoacyl-tRNA hyd
16	66	16.8	303	2 S60549	envelope polyprote
17	66	16.8	303	2 S60550	envelope polyprote
18	65.5	16.6	504	2 T52254	cytochrome P450 [1
19	65.5	16.6	797	2 D70581	hypothetical prote
20	65	16.5	333	2 B64380	hypothetical prote
21	64.5	16.4	165	2 T12774	hypothetical yolk
22	64.5	16.4	490	2 B86265	cytochrome P450 71
23	64	16.2	838	2 T47828	hypothetical prote
24	63.5	16.1	374	2 T02315	hypothetical prote
25	62	15.7	297	2 S60538	envelope polyprote
26	62	15.7	362	2 T34432	hypothetical prote
27	62	15.7	527	2 S42512	recombination-acti
28	62	15.7	662	2 S55274	alpha-N-arabino-fur
29	61.5	15.6	276	2 H70869	probable Enoyl-CoA

30	61	15.5	71	2 S42911	env polypeptide -
31	61	15.5	177	1 WZVZA9	20K HindIII-C prot
32	61	15.5	299	2 S60523	envelope polyprote
33	61	15.5	575	2 P69432	fumaryl reductase
34	60.5	15.4	339	1 LUCH2	annexin II - chick
35	60.5	15.4	346	2 S55403	cyma protein precu
36	60.5	15.4	1275	2 A38985	nucleotide exchang
37	60	15.2	69	2 A25161	tray protein - Sal
38	60	15.2	71	2 S42901	env polypeptide -
39	60	15.2	71	2 S42908	env polypeptide -
40	60	15.2	71	2 S42898	env polypeptide -
41	60	15.2	71	2 S42912	env polypeptide -
42	60	15.2	84	2 S76443	hypothetical prote
43	60	15.2	214	2 H82745	acyl-[ACPL]-UDP-N-a
44	60	15.2	504	2 D95985	probable sugar upt
45	60	15.2	563	2 D81134	D-lactate dehydrog

ALIGNMENTS

RESULT 1

T25763  
hypothetical protein F46F11.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25763  
R:Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, February 1997  
A:Description: The sequence of C. elegans cosmid F46F11.  
A:Reference number: Z20083  
A:Accession: T25763  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-73 <PAU>  
A:Cross-references: UNIPROT:P91302; UNIPARC:UPI000007680C; EMBL:U88173; PIDN:AAB42266.1;  
A:Experimental source: strain Bristol N2; clone F46F11  
C:Genetics:  
A:Gene: CESP:F46F11.4  
A:Map position: 1  
A:Introns: 38/2

Query Match	84.0%	Score	331	DB	2	Length	73
Best Local Similarity	80.8%	Pred. No.	1.4e-30				
Matches	59	Conservative	6	Mismatches	8	Indels	0
Gaps	0						
Qy	1	MIEVVCNDRGLGKVRKVCNTDDTIGDLKLI	AAQTGRWNKIVLKKWYTFKDHVSLG	60			
Db	1	MIEITVNDRLGKVRKVCNPSDTIGDLKLI	AAQTGRWEKIVLKKWYTYKDHITLMDY	60			
Qy	61	EIH DGMNLELYYQ	73				
Db	61	EIH EGFNFELYYQ	73				

RESULT 2

T47458  
hypothetical protein T14D3.120 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47458  
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24467  
A:Accession: T47458  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <JOR>  
A:Cross-references: UNIPROT:Q9M1U1; UNIPARC:UPI00000A81D7; EMBL:AL138649  
A:Experimental source: cultivar Columbia; BAC clone T14D3  
C:Genetics:  
A:Map position: 3  
A>Note: T14D3.120

```
Query Match      77.7%; Score 306; DB 2; Length 73;
Best Local Similarity 79.2%; Pred. No. 9.4e-28;
Matches 57; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGRWNKIVLKKWYTFKDHVSLGDY 60
DB 1 MIEVLNDRLGKKVRVKCNEDTIGDLKKLVAAQTGTREKIRIQKNYIKDHIPLKDY 60

QY 61 EIHGGMNLELYY 72
DB 61 EIHGGMGLELYY 72

RESULT 3
T40200
ubiquitin-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40200
R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21912
A:Accession: T40200
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-73 <OLI>
A:Cross-references: UNIPROT:O94650; UNIPARC:UPI000006A678; EMBL:AL049190; PIDN:CAB39137.
A:Experimental source: strain 972h-; cosmid c31E1
C:Genetics:
A:Gene: SPDB:SPBC31E1.03
A:Map position: 2
A:Introns: 6/2; 40/2

Query Match      76.4%; Score 301; DB 2; Length 73;
Best Local Similarity 73.6%; Pred. No. 3.5e-27;
Matches 53; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGRWNKIVLKKWYTFKDHVSLGDY 60
DB 1 MIEVLNDRLGKKVRVKCMPDDTIGDFFKLVAAQTGTPRRIVLKKWHSVFKDNITLADY 60

QY 61 EIHGGMNLELYY 72
DB 61 EIHGGMSELYY 72

RESULT 4
S78735
protein YNR032c-a - Yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 15-Jan-1999
C:Accession: S78735
R:Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63346
A:Accession: S78735
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-73 <POH>
A:Cross-references: UNIPARC:UPI0000168403; EMBL:Z71647; MIPS:YNR032c-a
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 14R

Query Match      64.0%; Score 252; DB 2; Length 73;
Best Local Similarity 65.3%; Pred. No. 1.3e-21;
Matches 47; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGRWNKIVLKKWYTFKDHVSLGDY 60
DB 1 MIEVWVNDRLGKKVRVKCLAEDSVGDFKVLSLQIGTPQPNKIVLKGGSVLKDHISLEDY 60
```

```
QY 61 EIHGGMNLELYY 72
DB 61 EVHDQTNLELYY 72

RESULT 5
S55243
ubiquitin-like protein 8 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Oct-2004
C:Accession: S55243; S61068
R:Callis, J.; Carpenter, T.; Sun, C.W.; Vierstra, R.D.
Genetics 139, 921-939, 1995
A:Title: Structure and evolution of genes encoding polyubiquitin and ubiquitin-like prot
A:Reference number: S55243; MUID:95229071; PMID:7713442
A:Accession: S55243
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-631 <CAL>
A:Cross-references: UNIPROT:Q39256; UNIPARC:UPI000017A48A; EMBL:L05917
A:Experimental source: ecotype Columbia
R:Callis, J.; Carpenter, T.; Sun, C.W.; Vierstra, R.D.
submitted to the EMBL Data Library, June 1995
A:Reference number: S61068
A:Accession: S61068
A:Molecule type: DNA
A:Residues: 1-341, E', 343-631 <CAN>
A:Cross-references: UNIPARC:UPI00000A3588; EMBL:L05917; NID:9870793; PIDN:AAA68879.1; PII
F:79-154/Domain: ubiquitin homology <UBH1>
F:155-237/Domain: ubiquitin homology <UBH2>
F:238-318/Domain: ubiquitin homology <UBH3>
F:319-392/Domain: ubiquitin homology <UBH4>
F:393-468/Domain: ubiquitin homology <UBH5>
F:469-551/Domain: ubiquitin homology <UBH6>
F:552-627/Domain: ubiquitin homology <UBH8>

Query Match      20.3%; Score 80; DB 2; Length 631;
Best Local Similarity 31.7%; Pred. No. 0.41;
Matches 19; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY 11 GKKVRVKCNTDTIGDLKKLIAAQTGRWNKIVLKKWYTFKDHVSLGDIYHIDGMNLEL 70
DB 561 GXTIILEVSSDTIANVKEIKVKGKPDQQLIFFGQQLDGVTLGDYDIHKKSTLYL 620

RESULT 6
B95037
hyaluronidase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95037
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95037
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1066 <KUR>
A:Cross-references: UNIPROT:Q54873; UNIPARC:UPI000012D052; GB:AE005672; PIDN:AAK74491.1.
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0314

Query Match      18.9%; Score 74.5; DB 2; Length 1066;
Best Local Similarity 34.3%; Pred. No. 3;
Matches 23; Conservative 8; Mismatches 27; Indels 9; Gaps 3;

QY 13 KVRVKCNTDDTIGDLKKLIAAQTG---TRWNKIV---LKKWYTFKDHVSLGDIYHIDGM 66
```



```
Db      128 KLRFKIKTDNKIGIAKVIIEESGDKRLWNSATTSGTKDQTIHEADYSPTLDV---DKI 184
Qy      67 NLELYYQ 73
Db      185 KLELFYE 191

RESULT 7
F97907
hyaluronate lyase (EC 4.2.2.1) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: F97907
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Accession: F97907
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1078 <KUR>
A:Cross-references: UNIPROT:Q8CWU3; UNIPARC:UPI000000E4928; GB:AE007317; PIDN:AAK99090.1;
C:Genetics:
A:Gene: hysA
C:Keywords: carbon-oxygen lyase

Query Match      18.7%; Score 73.5; DB 2; Length 1078;
Best Local Similarity 32.8%; Pred. No. 4;
Matches 22; Conservative 9; Mismatches 27; Indels 9; Gaps 3;

Qy      13 KVRVKCNTDDTIGDLKKLIAAQTG---TRWNKIV---LKKWYTFKDHVSLGDIYEHGDM 66
Db      140 KLRFKIKTDNKVGIAKVIIEESGDKRLWNSATTSGTKDQTIHEADYSPTLDV---DKI 196

Qy      67 NLELYYQ 73
Db      197 KLELFYE 203

RESULT 8
S60529
envelope polyprotein gp41 - human immunodeficiency virus type 1 (isolate CI-14-13) (frac
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate CI-14-13
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60529
R:Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Franssen, K.; Motte, J.;
AIDS 8, 21-26, 1994
A:Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d
A:Reference number: S60521; MUID:94280700; PMID:8011235
A:Accession: S60529
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <JAN>
A:Cross-references: UNIPROT:Q76172; UNIPARC:UPI0000010385B; EMBL:X72031; NID:9468637; PID
A:Experimental source: isolate CI-14-13
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; glycoprotein; immunodeficiency; polyprotein

Query Match      17.9%; Score 70.5; DB 2; Length 299;
Best Local Similarity 27.7%; Pred. No. 2.2;
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 3;

Qy      9 RLGKVRVKCN-----TDDTIGDLKKLIAAQTGTRWNKIV----- 43
Db      20 QLVKPVIRNTRPSNTRKSVIPGPGAFATDIIIGIRQAHNCVSKTEWNETLQOVAT 79

RESULT 9
S60528
envelope polyprotein gp41 - human immunodeficiency virus type 1 (isolate CI-14-5 and iso
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate CI-14-5; isolate CI-14-21
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60528; S60530
R:Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Franssen, K.; Motte, J.;
AIDS 8, 21-26, 1994
A:Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d
A:Reference number: S60521; MUID:94280700; PMID:8011235
A:Accession: S60528
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <JAN>
A:Cross-references: UNIPROT:Q76171; UNIPROT:Q76173; UNIPARC:UPI0000101A6F; EMBL:X72030;
A:Experimental source: isolate CI-14-5
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
A:Accession: S60530
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <JAN>
A:Cross-references: UNIPARC:UPI0000101A6F; EMBL:X72032; NID:9468639; PIDN:CAA50915.1; P1
A:Experimental source: isolate CI-14-21
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; glycoprotein; immunodeficiency; polyprotein

Query Match      17.9%; Score 70.5; DB 2; Length 299;
Best Local Similarity 27.7%; Pred. No. 2.2;
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 3;

Qy      9 RLGKVRVKCN-----TDDTIGDLKKLIAAQTGTRWNKIV----- 43
Db      20 QLVKPVIRNTRPSNTRKSVIPGPGAFATDIIIGIRQAHNCVSKTEWNETLQOVAT 79

RESULT 10
T29404
hypothetical protein C16C8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29404
R:Waterston, R.; Le, T.T.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans cosmid C16C8.
A:Reference number: Z20617
A:Accession: T29404
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WAT>
A:Cross-references: UNIPROT:P91057; UNIPARC:UPI000007F11D; EMBL:U80452; PIDN:AAB37860.1;
A:Experimental source: strain Bristol N2; clone C16C8
C:Genetics:
A:Gene: CRSP:C16C8.4
A:Map position: 2
A:Introns: 27/3; 199/1

Query Match      17.5%; Score 69; DB 2; Length 226;
Best Local Similarity 26.5%; Pred. No. 2.4;
Matches 18; Conservative 13; Mismatches 29; Indels 8; Gaps 2;
```

```
QY 10 LGKKVRVKNCTDDTIGDLKLLIAAQ-----TGTRWNKIVLKKWYTIFKDHVSLGDDYIEHDG 65
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 LGVSYAFKIHREDTVDIKNDIEHRHDIPQHSYWLFSFGKR-----LEDHCSIGDYNIQKS 216
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 MNLELYYQ 73
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 SITMYFR 224
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
A:Accession: A70363
A:Title: mannose-1-phosphate guanylttransferase - Aquifex aeolicus
A:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: A70363
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: A70363
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-831 <AQF>
A:Cross-references: UNIPROT:O66933; UNIPARC:UPI000005641F; GB:AE000704; NID:g2983301; P
A:Experimental source: strain VF5
C:Genetics:
A:Gene: mpg

Query Match 17.4%; Score 68.5; DB 2; Length 831;
Best Local Similarity 27.4%; Pred. No. 11;
Matches 23; Conservative 17; Mismatches 21; Indels 23; Gaps 5;

QY 4 VVCND-RLGKKVRVK-----CNTDDTIGDLKLLIAAQGTGRWNKIVLKKWYTIFKDH 54
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 322 VICNDVKGKRVKAKGWTAECDCEDEVLFKDVVV-----WPEKIEKGSVVTKNI 375
: : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 VSLGDYE--IH DG-----MNLEL 70
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 VCESKWEKGIFGNKVINVEL 399
: : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
S60545
envelope polyprotein gp41 - human immunodeficiency virus type 1 (isolate CI-45-1) (fragme
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate CI-45-1
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60545
R:Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Fransen, K.; Motte, J.;
AIDS 8, 21-26, 1994
A:Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d
A:Reference number: S60521; MUID:94280700; PMID:8011235
A:Accession: S60545
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-294 <JAN>
A:Cross-references: UNIPROT:Q76190; UNIPARC:UPI0000101898; EMBL:X72047; NID:g468669; PID
A:Experimental source: isolate CI-45-1
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; Glycoprotein; immunodeficiency; polyprotein

Query Match 17.3%; Score 68; DB 2; Length 294;
Best Local Similarity 26.8%; Pred. No. 4.2;
Matches 22; Conservative 11; Mismatches 21; Indels 28; Gaps 3;

QY 9 RLGKKVRVKCN-----TDDTIGDLKLLIAAQGTGRWNKIV----- 43
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 QLAKPVKINCTRPNNNTRKSVRIGPGQTFYATGDIIGDIRQAHCVNSGTWNETLQKVA 79
: : : : : : : : : : : : : : : : : : : : : : : : :

QY 54 VVCND-RLGKKVRVK-----CNTDDTIGDLKLLIAAQGTGRWNKIVLKKWYTIFKDH 54
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 322 VICNDVKGKRVKAKGWTAECDCEDEVLFKDVVV-----WPEKIEKGSVVTKNI 375
: : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 VSLGDYE--IH DG-----MNLEL 70
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 VCESKWEKGIFGNKVINVEL 399
: : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
S60524
envelope polyprotein gp41 - human immunodeficiency virus type 1 (isolate CI-45-3) (fragme
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate CI-45-3
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60524
R:Janssens, W.; Heyndrickx, L.; Van de Peer, Y.; Bouckaert, A.; Fransen, K.; Motte, J.;
AIDS 8, 21-26, 1994
A:Title: Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Cote d
A:Reference number: S60521; MUID:94280700; PMID:8011235
A:Accession: S60524
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-294 <JAN>
A:Cross-references: UNIPROT:Q76168; UNIPARC:UPI000010200C; EMBL:X72027; NID:g468780; PID
A:Experimental source: isolate CI-45-3
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; Glycoprotein; immunodeficiency; polyprotein

Query Match 17.3%; Score 68; DB 2; Length 294;
Best Local Similarity 26.8%; Pred. No. 4.2;
Matches 22; Conservative 11; Mismatches 21; Indels 28; Gaps 3;

QY 9 RLGKKVRVKCN-----TDDTIGDLKLLIAAQGTGRWNKIV----- 43
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 QLAKPVKINCTRPNNNTRKSVRIGPGQTFYATGDIIGDIRQAHCVNSGTWNETLQKVA 79
: : : : : : : : : : : : : : : : : : : : : : : : :

QY 44 -LKKWY--TIFKDHVSLGDDYEI 62
||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 QLRKHFNKTIIFASPSGGDVEI 101
: : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
G96610
probable disease resistance protein RPPI-waA [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G96610
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luroso, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-709 <STO>
A:Cross-references: UNIPROT:Q9FV77; UNIPARC:UPI000009DF9E; GB:AE005173; NID:g11055850; P
C:Genetics:
A:Gene: T8L23.12
A:Map position: 1

Query Match 17.3%; Score 68; DB 2; Length 709;
Best Local Similarity 26.4%; Pred. No. 11;
Matches 14; Conservative 13; Mismatches 22; Indels 4; Gaps 2;

QY 21 DDTIGDLKK-LIAAQGTGRWNKIVLKKWYTIFKDHVSLGDDYIEHDGMNLELY 72
||| : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: April 8, 2006, 00:04:11  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 23:56:33 ; Search time 233 Seconds  
(without alignments)  
221.045 Million cell updates/sec

Title: US-10-067-832D-2

Perfect score: 394

Sequence: 1 MEVVCNDRLGKKVRVKNT.....HVSGLDGYEIHGNNLELYQ 73

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	394	100.0	73	1	UBL5 HUMAN
2	394	100.0	73	1	UBL5_MESAU
3	394	100.0	73	1	UBL5 MOUSE
4	394	100.0	73	1	UBL5_PSAOB
5	394	100.0	73	2	Q5RC10_PONPY
6	394	100.0	73	2	Q4RSJ3_MACPA
7	368	93.4	73	1	UBL5_BEARE
8	367	93.1	73	2	Q5H2C1_XENLA
9	354	89.8	73	2	Q4SMI6_TETNG
10	344	87.3	73	2	Q7QHM2_ANOGA
11	339	86.0	73	1	UBL5_DROME
12	331	84.0	73	1	UBL5_CAEEL
13	331	84.0	73	2	Q617M9_CAEBR
14	329	83.5	73	2	Q5BS57_SCHUA
15	325	82.5	73	2	Q4VAF8_MOUSE
16	315	79.9	73	2	Q6K220_ORYSA
17	313	79.4	73	1	UBL5_ARATH
18	313	79.4	73	2	Q570V8_ARATH
19	308	78.2	73	2	Q7R8W4_PLAYO
20	308	78.2	73	2	Q4YUJ7_PLABTH
21	306	77.7	73	2	Q9MIU1_ARATH
22	302	76.6	73	2	Q81545_PLAF7
23	301	76.4	73	1	HUB1_SCHPO
24	293	74.4	73	2	Q4M259_THERA
25	291	73.9	76	2	Q4UBE9_THEAN
26	288	73.1	276	2	Q5SM54_CRYNE
27	288	73.1	276	2	Q5K8L5_CRYNE
28	266	67.5	73	1	HUB1_DBPHA
29	263	66.8	70	2	Q6C104_YARLI
30	250	63.5	47	2	Q5BJL1_RAT
31	248	62.9	73	1	HUB1_YEAST

32	243	61.7	73	1	HUB1_CANGA
33	238.5	60.5	74	1	HUB1_KLULA
34	238.5	60.5	79	2	Q6CUI1_KLULA
35	235	59.6	73	1	HUB1_ASHGO
36	231	58.6	87	2	Q54Q03_DICTDI
37	214	54.3	79	2	Q50903_ENTHI
38	168	42.6	73	2	Q4P9W2_USTWA
39	154.5	39.2	237	2	Q4I016_GIBZE
40	141.5	35.9	616	2	Q5BCG2_EMENI
41	123.5	31.3	239	2	Q526Z0_MAGGR
42	119.5	30.3	261	2	Q8X065_NEUCR
43	113.5	28.8	189	2	Q4WYK6_ASPTU
44	94	23.9	219	2	Q6DWES_9HIVI
45	93	23.6	115	2	Q7ZCU2_9HIVI

#### ALIGNMENTS

RESULT 1  
ID UBL5\_HUMAN STANDARD; PRT; 73 AA.  
AC Q9BZL1;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DE 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Ubiquitin-like protein 5.  
GN Names=UBL5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RC TISSUE=iris;  
RX MEDLINE=21100898; PubMed=11161819; DOI=10.1006/geno.2000.6439;  
RA Friedman J.S., Koop B.F., Raymond V., Walter M.A.;  
RT "Isolation of a ubiquitin-like (UBL5) gene from a screen identifying  
RT highly expressed and conserved iris genes.";  
RL Genomics 71:252-255(2001).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RA Kainane N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
RA Phelan M., Farmer A.;  
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor  
RT vector.";  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uudin T.B., Toohilyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Q6fix7	candida gla
Q6cul2	kluyveromyc
Q6cul1	kluyveromyc
Q756x3	ashbya gose
Q54q03	dictyoateli
Q50p03	entamoeba h
Q4p9w2	utellago ma
Q4i016	gibbarella
Q5bcg2	aspergillus
Q526z0	magnaporthe
Q8x065	neurospora
Q4wyk6	aspergillus
Q6dwe5	human immun
Q7zcu2	human immun



60,770 full-length cDNAs.";  
 Nature 420:563-573 (2002).  
 [2]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 TISSUE=Mammary gland;  
 MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heide F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Ueudin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Vallalath D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences".  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- SUBUNIT: Interacts with CLK4 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AK002730; BAB23112.1; -; mRNA.  
 DR EMBL; AK003592; BAB23111.1; -; mRNA.  
 DR EMBL; AK007726; BAB25215.1; -; mRNA.  
 DR EMBL; AK009854; BAB26545.1; -; mRNA.  
 DR EMBL; AK012803; BAB28481.1; -; mRNA.  
 DR EMBL; AK051149; BAC34537.1; -; mRNA.  
 DR EMBL; BC028498; AAB28498.1; -; mRNA.  
 DR PUB; LUH6; NMR; A=1-73.  
 DR MGI; MGI:1913427; Ub1.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50053; UBIQUITIN\_2; FALSE\_NEG.  
 KW 3D-structure; Ub1 conjugation pathway.  
 FT DOMAIN 1 73 Ubiquitin-like.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;  
 Query Match 100.0%; Score 394; DB 1; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-37;  
 Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MIEVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGTGRNWKIVLKKWYTFKDHVSLGDY 60  
 DB 1 MIEVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGTGRNWKIVLKKWYTFKDHVSLGDY 60  
 QY 61 EIHGGMNLELYQ 73  
 DB 61 EIHGGMNLELYQ 73  
 RESULT 5  
 QSRC10\_PONPY PRELIMINARY; PRT; 73 AA.  
 AC QSRC10;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DE Hypothetical protein DKFZp469G145.  
 GN Name=DKFZp469G145;  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
 OC Pongo.  
 OC NCBI\_TaxID=9600;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC The German cDNA Consortium;  
 RA Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
 RA Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR58290; CAH90527.1; -; mRNA.  
 DR SMR; QSRC10; 1-73.  
 DR GO; GO:0006464; P:protein modification; IEA.

Ubiquitin-like protein 5 (Beacon protein).  
 Name=UBL5;  
 OS Psammomys obesus (Fat sand rat).  
 OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Gerbillinae; Psammomys.  
 OC NCBI\_TaxID=48139;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Hypothalamus;  
 MEDLINE=20527879; PubMed=11078442;  
 Collier G.R., McMillan J.S., Windmill K., Walder K., Tenne-Brown J.,  
 de Silva A., Trevaaskis J., Jones S., Morton G.J., Lee S., Augert G.,  
 Civitarese A., Zimet P.Z.;  
 "Beacon: a novel gene involved in the regulation of energy balance.";  
 Diabetes 49:1766-1771 (2000).  
 RL SUBUNIT: Interacts with CLK4 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AF318186; AAG34704.1; -; mRNA.  
 DR SMR; Q791B0; 1-73.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50053; UBIQUITIN\_2; FALSE\_NEG.  
 KW Ub1 conjugation pathway.  
 FT DOMAIN 1 73 Ubiquitin-like.  
 SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;  
 Query Match 100.0%; Score 394; DB 1; Length 73;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-37;  
 Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MIEVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGTGRNWKIVLKKWYTFKDHVSLGDY 60  
 DB 1 MIEVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGTGRNWKIVLKKWYTFKDHVSLGDY 60  
 QY 61 EIHGGMNLELYQ 73  
 DB 61 EIHGGMNLELYQ 73  
 RESULT 5  
 QSRC10\_PONPY PRELIMINARY; PRT; 73 AA.  
 AC QSRC10;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DE Hypothetical protein DKFZp469G145.  
 GN Name=DKFZp469G145;  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
 OC Pongo.  
 OC NCBI\_TaxID=9600;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC The German cDNA Consortium;  
 RA Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
 RA Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR58290; CAH90527.1; -; mRNA.  
 DR SMR; QSRC10; 1-73.  
 DR GO; GO:0006464; P:protein modification; IEA.

```

DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Query Match 100.0%; Score 394; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
DB 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
QY 61 EIHGGMNLELYQ 73
DB 61 EIHGGMNLELYQ 73

RESULT 6
Q4RSJ3 MACFA
ID Q4RSJ3_MACFA PRELIMINARY; PRT; 73 AA.
AC Q4RSJ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Brain cDNA, clone: OFIA-12296, similar to human ubiquitin-like 5
DE (UBL5), (Testis cDNA, clone: QcA-15893, similar to human ubiquitin-
DE like 5 (UBL5)).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB169550; BAE01632.1; -; mRNA.
DR EMBL; AB168933; BAE01034.1; -; mRNA.
SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Query Match 100.0%; Score 394; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
DB 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
QY 61 EIHGGMNLELYQ 73
DB 61 EIHGGMNLELYQ 73

RESULT 7
UBL5 BRARE
ID UBL5 BRARE STANDARD; PRT; 73 AA.
AC Q7SXF2;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ubiquitin-like protein 5.
GN Namesub15; ORFNames=zgc:66388;
OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA).
RG NIH - Zebrafish Gene Collection (ZGC) project;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BC055630; AAH55630.1; -; mRNA.
DR SMR; Q7SXF2; 1-72.
DR ZFIN; ZDB-GENE-040426-1629; zgc:66388.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PROSITE; PS00299; UBIQUITIN_1; FALSE_NEG.
DR PROSITE; PS0053; UBIQUITIN_2; FALSE_NEG.
KW Ub1 conjugation pathway.
FT DOMAIN 1 73 Ubiquitin-like.
SQ SEQUENCE 73 AA; 8587 MW; 23716CB90FC7C545 CRC64;

Query Match 93.4%; Score 368; DB 1; Length 73;
Best Local Similarity 94.4%; Pred. No. 1.9e-34;
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
DB 1 MIEVNCNDRLGKVRVKCNKNTDITIGDLKKLIAAQTGRNKKVILKKWYTFKDHVSLG DY 60
QY 61 EIHGGMNLELYQ 72
DB 61 EIHGGMNLELYQ 72

RESULT 8
Q5HZC1 XENLA
ID Q5HZC1_XENLA PRELIMINARY; PRT; 73 AA.
AC Q5HZC1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauener R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore J., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```



RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC089084; AAH89084.1; -; mRNA.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; Ubiquitin; 1.  
 DR PROSITE; PS0053; UBIQUITIN\_2; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 73 AA; 8635 MW; 2A51E123F36FF390 CRC64;  
  
 Query Match 93.1%; Score 367; DB 2; Length 73;  
 Best Local Similarity 91.8%; Pred. No. 2.46-34;  
 Matches 67; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
 Qy 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
 Db 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
  
 Qy 61 EIHGGMNLELYQ 73  
 Db 61 EIHGGMNLELYQ 73  
  
 RESULT 9  
 ID Q4SM16\_TETNG PRELIMINARY; PRT; 73 AA.  
 AC Q4SM16\_TETNG  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Chromosome 18 SCAF14547, whole genome shotgun sequence.  
 DE (Fragment).  
 GN ORFNames=GSTENG00015743001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,  
 RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,  
 RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,  
 RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype";  
 RL Nature 431:946-957 (2004).

RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAB01014547; CAP98146.1; -; Genomic\_DNA.  
 FT NON TER 73  
 SQ SEQUENCE 73 AA; 8592 MW; 6A14CA0323F8E8EB CRC64;  
  
 Query Match 89.8%; Score 354; DB 2; Length 73;  
 Best Local Similarity 90.4%; Pred. No. 7.66-33;  
 Matches 66; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
  
 Qy 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
 Db 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
  
 Qy 61 EIHGGMNLELYQ 73  
 Db 61 EIHGGMNLELYQ 73  
  
 RESULT 10  
 ID Q7QHM2\_ANOGA PRELIMINARY; PRT; 73 AA.  
 AC Q7QHM2;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE ENSANGP0000015674.  
 GN ORFNames=ENSANG0000013185;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;  
 OC Anophelinae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation.";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008816; EAA05206.2; -; Genomic\_DNA.  
 DR SNR; Q7QHM2; 1-73.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; Ubiquitin; 1.  
 DR PROSITE; PS0053; UBIQUITIN\_2; 1.  
 SQ SEQUENCE 73 AA; 8599 MW; 43D212BE601B3C0D CRC64;  
  
 Query Match 87.3%; Score 344; DB 2; Length 73;  
 Best Local Similarity 83.6%; Pred. No. 1.1e-31;  
 Matches 61; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
  
 Qy 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
 Db 1 MIEVVCNDRLGKVRKVCNTDDTTIGDLKLLIAAQTGRNKNVILKKWYTFKDHVSLG DY 60  
  
 Qy 61 EIHGGMNLELYQ 73  
 Db 61 EIHGGMNLELYQ 73

```

RESULT 11
UBLS DROME                                STANDARD;          PRT;          73 AA.
AC Q9V9S8; Q4QPS6;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ubiquitin-like protein 5.
GN Name1(2)k03203; ORFNames=CG3450;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]_
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Baldwin D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobaric C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasatman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Mirza S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Rettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=Berkelley;
RA Stapleton M., Carlson J.W., Chavez C., Frise E., George R.A.,

```

```

RA Pacle J.M., Park S., Wan K.H., Yu C., Celniker S.E.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB003790; AAF57398.1; -; Genomic_DNA.
CC EMBL; BT023690; AAY85090.1; -; mRNA.
CC SMR; Q9V9S8; 1-73.
CC Ensembl; CG3450; Drosophila melanogaster.
CC FlyBase; FBgn0022224; l(2)k03203.
CC GO; GO:0005737; Cytoplasm; ISS.
CC InterPro; IPR000626; Ubiquitin.
CC Pfam; PF00240; ubiquitin; 1.
CC PROSITE; PS00299; UBIQUITIN_1; FALSE_NEG.
CC PROSITE; PS00053; UBIQUITIN_2; 1.
KW Ubiquitin-like.
FT DOMAIN 1 73 Ubiquitin-like.
SQ SEQUENCE 73 AA; 8570 MW; 0870111AC5686A70 CRC64;
Query Match 86.0%; Score 339; DB 1; Length 73;
Best Local Similarity 86.3%; Pred. No. 4e-31; Indels 0; Gaps 0;
Matches 63; Conservative 3; Mismatches 0;
QY 1 MIEVNCNDRLGKVRKVCNDDTIGDLKKLIAAQTGTRNKKVLFKQWYTFKDHVSLGDY 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 EIHGGMNLELYQ 73
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 EIHGGMNLELYQ 73
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 12
UBLS_CAEEL STANDARD; PRT; 73 AA.
ID UBLS_CAEEL
AC F91302;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ubiquitin-like protein 5.
GN NameSubl-5; ORFNames=F46F11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U88173; AAK21382.1; -; Genomic_DNA.
CC F91302; T25763.
CC SMR; F91302; 1-73.
CC Ensembl; F46F11.4; Caenorhabditis elegans.
CC WormBase; WBGen00006726; ubl-5.

```

```
DR WormPep; F46F11.4; CE10602.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF0240; ubiquitin; 1.
DR PROSITE; PS00299; UBIQUITIN 1; FALSE_NEG.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
DR Complete proteome; Ub1 conjugation pathway.
FT DOMAIN 1 Ubiquitin-like.
SQ SEQUENCE 73 AA; 8738 MW; 61CA839BBA4006A4 CRC64;

Query Match 84.0%; Score 331; DB 1; Length 73;
Best Local Similarity 80.8%; Pred. No. 3.3e-30;
Matches 59; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRLGKVRKCNPTDITGDLKLLIAAQTGRWKNVILKKWYTFKDHVSLG DY 60
Db 1 MIEITVNDRLGKVRKCNPSDITGDLKLLIAAQTGRWEKIVLKKWYTFKDHITLMDY 60
Qy 61 EIHGGMNLELYYQ 73
Db 61 EIHEGFNFELYQ 73

RESULT 13
Q617M9 CAEBR PRELIMINARY; PRT; 73 AA.
AC Q617M9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein CBG14905.
GN Name=CBG14905;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; CAAC01000608; CAB68926.1; -; Genomic_DNA.
DR SRR; Q617M9; 1-73.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF0240; ubiquitin; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8738 MW; 61CA839BBA4006A4 CRC64;

Query Match 84.0%; Score 331; DB 2; Length 73;
Best Local Similarity 80.8%; Pred. No. 3.3e-30;
Matches 59; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRLGKVRKCNPTDITGDLKLLIAAQTGRWKNVILKKWYTFKDHVSLG DY 60
Db 1 MIEITVNDRLGKVRKCNPSDITGDLKLLIAAQTGRWEKIVLKKWYTFKDHITLMDY 60
Qy 61 EIHGGMNLELYYQ 73
Db 61 EIHEGFNFELYQ 73

RESULT 14
Q5BS57 SCHJA PRELIMINARY; PRT; 73 AA.
AC Q5BS57;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Mammary gland;
RC NTH MGC Project;
RG Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC096400; AAH96400.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8582 MW; 9B2C997752A6068 CRC64;

Query Match 82.5%; Score 325; DB 2; Length 73;
Best Local Similarity 83.6%; Pred. No. 1.6e-29;
Matches 61; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

RN NUCLEOTIDE SEQUENCE.
RP Han Z.;
RL EMBL; AY915407; AAX30628.1; -; mRNA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF0240; ubiquitin; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8679 MW; D3A862289A0B7CB5 CRC64;

Query Match 83.5%; Score 329; DB 2; Length 73;
Best Local Similarity 79.5%; Pred. No. 5.7e-30;
Matches 58; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRLGKVRKCNPTDITGDLKLLIAAQTGRWKNVILKKWYTFKDHVSLG DY 60
Db 1 MIEITCNDRLGKVRKCNPTDKVSDLKLLIAAQTGRWKNVILKKWYTFKDHITLQDY 60
Qy 61 EIHGGMNLELYYQ 73
Db 61 EINDGMNLELYYQ 73

RESULT 15
Q4VAF8 MOUSE PRELIMINARY; PRT; 73 AA.
AC Q4VAF8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG STRAIN=C57BL/6J; TISSUE=Mammary gland;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RG NTH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC096400; AAH96400.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8582 MW; 9B2C997752A6068 CRC64;

Query Match 82.5%; Score 325; DB 2; Length 73;
Best Local Similarity 83.6%; Pred. No. 1.6e-29;
Matches 61; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
```

Qy 1 MIEVVCNDRLGKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTFPKDHVSLG DY 60  
Db | | | | | | | | | | | | | | | | | | | | | | : | | | | | | | |  
1 MTEWCNHYHLGKRVKVKYNTDDTIGDLKKLIEAQTGTHWNKILKKWYTFLODHVSLG DY 60  
Qy 61 EIHGNNLELYQ 73  
Db | | | | | | | | | | | | | | | | | | | | | | : | | | | | | | |  
61 EIHGNNLELSYQ 73

Search completed: April 8, 2006, 00:03:25  
Job time : 233 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2006, 00:03:43 ; Search time 47 Seconds  
(without alignments)  
128.411 Million cell updates/sec

Title: US-10-067-832D-2  
Perfect score: 394  
Sequence: 1 MIEVVCNDRLGKVRKVCNT.....HVSIGDYEHHDGMNLELYQ 73

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	100.0	73	2	US-09-331-930A-2
2	394	100.0	73	2	US-09-331-930A-19
3	394	100.0	73	2	US-09-331-930A-20
4	394	100.0	73	2	US-09-331-930A-21
5	370	93.9	73	2	US-09-331-930A-24
6	354	89.8	73	2	US-09-331-930A-25
7	331	84.0	73	2	US-09-331-930A-22
8	286	72.6	66	2	US-09-331-930A-23
9	252	64.0	75	2	US-09-248-796A-16302
10	182	46.2	33	2	US-09-331-930A-16
11	182	46.2	33	2	US-09-331-930A-18
12	163	41.4	33	2	US-09-331-930A-14
13	80	20.3	60	2	US-09-331-930A-7
14	73.5	18.7	1078	2	US-09-583-110-4036
15	73.5	18.7	1080	2	US-09-107-433-4843
16	69	17.5	535	2	US-09-107-532A-6214
17	69	17.5	860	2	US-10-290-579A-252
18	66	16.8	921	2	US-09-302-540-15467
19	63.5	16.1	467	2	US-10-040-802-8
20	63.5	16.1	846	2	US-10-290-579A-249
21	63	16.0	374	2	US-09-248-796A-16046
22	63	16.0	533	2	US-09-248-796A-20066
23	61.5	15.6	842	2	US-10-290-579A-254
24	61	15.5	867	2	US-09-540-236-3193
25	60	15.2	475	2	US-09-248-796A-20067
26	60	15.2	482	2	US-09-270-767-43292
27	60	15.2	494	2	US-08-889-841B-19

28	60	15.2	494	2	US-09-419-362-19	Sequence 19, Appl
29	60	15.2	556	2	US-09-248-796A-18763	Sequence 18763, A
30	59.5	15.1	415	2	US-09-248-796A-14232	Sequence 14232, A
31	59.5	15.1	475	2	US-09-949-016-7954	Sequence 7954, Ap
32	59.5	15.1	493	2	US-09-411-628-10	Sequence 10, Appl
33	59.5	15.1	493	2	US-10-174-794-10	Sequence 10, Appl
34	59.5	15.1	629	2	US-09-107-532A-6656	Sequence 6656, Ap
35	59	15.0	157	2	US-09-248-796A-17780	Sequence 17780, A
36	59	15.0	440	2	US-09-540-236-2067	Sequence 2067, Ap
37	59	15.0	455	2	US-08-889-841B-46	Sequence 46, Appl
38	59	15.0	455	2	US-09-419-362-46	Sequence 46, Appl
39	59	15.0	850	1	US-08-448-603A-28	Sequence 28, Appl
40	59	15.0	850	2	US-09-134-075-28	Sequence 28, Appl
41	59	15.0	850	2	US-09-492-739-28	Sequence 28, Appl
42	59	15.0	850	2	US-09-966-931A-28	Sequence 28, Appl
43	58.5	14.8	353	2	US-09-248-796A-15778	Sequence 15778, A
44	58.5	14.8	498	2	US-08-889-841B-2	Sequence 2, Appl
45	58.5	14.8	498	2	US-08-889-841B-5	Sequence 5, Appl

## ALIGNMENTS

## RESULT 1

US-09-331-930A-2

; Sequence 2, Application US/09331930A

; Patent No. 6436670

; GENERAL INFORMATION:

; APPLICANT: ZIMMET, PAUL Z.

; APPLICANT: COLLIER, GREGORY

; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR

; FILE REFERENCE: 22975-20007.00

; CURRENT APPLICATION NUMBER: US/09/331.930A

; CURRENT FILING DATE: 1999-06-30

; PRIOR APPLICATION NUMBER: PCT/AU98/00902

; PRIOR FILING DATE: 1998-10-30

; PRIOR APPLICATION NUMBER: AU PP0117/97

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: AU PP0323/97

; PRIOR FILING DATE: 1997-11-11

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 2

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: Amino acid

; OTHER INFORMATION: sequence for beacon from unknown organism

US-09-331-930A-2

Query Match 100.0%; Score 394; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 4.8e-46;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKXIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60

Db 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKXIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60

Qy 61 EIHGGMNLELYQ 73

Db 61 EIHGGMNLELYQ 73

## RESULT 2

US-09-331-930A-19

; Sequence 19, Application US/09331930A

; Patent No. 6436670

; GENERAL INFORMATION:

; APPLICANT: ZIMMET, PAUL Z.

; APPLICANT: COLLIER, GREGORY

; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR

; FILE REFERENCE: 22975-20007.00

```
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; TYPE: PRT
; LENGTH: 73
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid
; OTHER INFORMATION: sequence for beacon from unknown organism
US-09-331-930A-19

Query Match      100.0%; Score 394; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 4.8e-46;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYYQ 73

RESULT 3
US-09-331-930A-20
; Sequence 20, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 73
; ORGANISM: Homo sapiens
US-09-331-930A-20

Query Match      100.0%; Score 394; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 4.8e-46;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYYQ 73

RESULT 4
US-09-331-930A-21
; Sequence 21, Application US/09331930A
```

```
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-331-930A-21

Query Match      100.0%; Score 394; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 4.8e-46;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYYQ 73

RESULT 5
US-09-331-930A-24
; Sequence 24, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-331-930A-24

Query Match      93.9%; Score 370; DB 2; Length 73;
Best Local Similarity 93.1%; Pred. No. 8.8e-43;
Matches 67; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVVCNDRLGKKVRVKCNTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTIKDHVSLGDY 60

QY 61 EIHGGMNLELYY 72
   ||||||||||||
Db 61 EIHGGMNLELYY 72
```

```

Db      61 EIHEGFNFELYYQ 73

RESULT 8
US-09-331-930A-23
; Sequence 23, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331.930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Fasiola hepatica
US-09-331-930A-23

Query Match      72.6%; Score 286; DB 2; Length 66;
Best Local Similarity 81.8%; Pred. No. 2.1e-31;
Matches 54; Conservative 5; Mismatches 7; Indels 0; Gaps 0

Qy      8 DRLGKVRVKCNTDDTIGDLKCLIAAQTGTRWNKIVLKKWYTFPKDHVSLGDIHDCMN 67
Db      1 DRLGKVRVKCNPTDKVGDULKCLIAAQTGAPERIVLKKWYTFPKDHVTLRDYEINDGMN 60

Qy      68 LELYYQ 73
Db      61 LELYYQ 66

RESULT 9
US-09-248-796A-16302
; Sequence 16302, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16302
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16302

Query Match      64.0%; Score 252; DB 2; Length 75;
Best Local Similarity 61.1%; Pred. No. 1e-26;
Matches 44; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Qy      1 MIEVVCNDRLGKVRVKCNTDDTIGDLKCLIAAQTGTRWNKIVLKKWYTFPKDHVSLGDI 60
Db      3 MIEIQANDRLGKVIKUKLCLETDTIGDVKKILGIQTGLEKILKGYQVYKHITLDDY 62

Qy      61 EIHDCMNLELYY 72

```

```
Db 63 EIHDFGNFELY 74

RESULT 10
US-09-331-930A-16
; Sequence 16, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; TYPE: PRT
; LENGTH: 33
; ORGANISM: Psammomys obesus
US-09-331-930A-16

Query Match 46.2%; Score 182; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 KIVLKKWYTFKDHVSLGDIYHIDGNLLEYQ 73
Db 1 KIVLKKWYTFKDHVSLGDIYHIDGNLLEYQ 33

RESULT 11
US-09-331-930A-18
; Sequence 18, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 33
; ORGANISM: Psammomys obesus
US-09-331-930A-18

Query Match 46.2%; Score 182; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 KIVLKKWYTFKDHVSLGDIYHIDGNLLEYQ 73
Db 1 KIVLKKWYTFKDHVSLGDIYHIDGNLLEYQ 33

RESULT 12
US-09-331-930A-14
; Sequence 14, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; TYPE: PRT
; LENGTH: 33
; ORGANISM: Homo sapiens
; NAME/KEY: MOD_RES
; LOCATION: (15)
; OTHER INFORMATION: variable amino acid
US-09-331-930A-14

Query Match 41.4%; Score 163; DB 2; Length 33;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAA 33
Db 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAA 33

RESULT 13
US-09-331-930A-27
; Sequence 27, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-331-930A-27

Query Match 20.3%; Score 80; DB 2; Length 60;
Best Local Similarity 31.7%; Pred. No. 0.002;
Matches 19; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY 11 GKRVKVCNCTDDTIGDLKKLIAAQGTGRWNKIVLKKWYTFKDHVSLGDIYHIDGNLLEYQ 70
Db 1 GKTILEVSSDTIANVKEIKVKEGKPDQWMLIFFGQQLDGVTLGDIYHIDGNLLEYQ 60

RESULT 14
US-09-583-110-4036
; Sequence 4036, Application US/09583110
```



```
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4036
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4036

Query Match      18.7%; Score 73.5; DB 2; Length 1078;
Best Local Similarity 32.8%; Pred. No. 0.61;
Matches 22; Conservative 9; Mismatches 27; Indels 9; Gaps 3;

Qy      13 KVRVKCNTDDTIGDLKLLIAAQTG---TRNNKIV---LKKWYTIKDHVSLGDIYHIDGM 66
Db      140 KLRFKIKTDNKVGIAKVRRIIEESGKRLWNSATTSKTKMQTIADYSPTLDV---DKI 198
Qy      67 NLELYYQ 73
Db      197 KLELFYE 203

RESULT 15
US-09-107-433-4843
; Sequence 4843, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS
; AND THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4843:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1080 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1080
; SEQUENCE DESCRIPTION: SEQ ID NO: 4843:
US-09-107-433-4843
```

```
Query Match      18.7%; Score 73.5; DB 2; Length 1080;
Best Local Similarity 32.8%; Pred. No. 0.61;
Matches 22; Conservative 9; Mismatches 27; Indels 9; Gaps 3;

Qy      13 KVRVKCNTDDTIGDLKLLIAAQTG---TRNNKIV---LKKWYTIKDHVSLGDIYHIDGM 66
Db      142 KLRFKIKTDNKVGIAKVRRIIEESGKRLWNSATTSKTKMQTIADYSPTLDV---DKI 198
Qy      67 NLELYYQ 73
Db      199 KLELFYE 205
```

```
Search completed: April 8, 2006, 00:05:03
Job time : 48 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2006, 00:04:23 ; Search time 166 Seconds  
(without alignments)  
183.744 Million cell updates/sec

Title: US-10-067-832D-2  
Perfect score: 394  
Sequence: 1 MIEVVCNDRLGKVRKVCNT.....HVSIGDYIHDGMNLELYQ 73

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main.\*  
1: /cgm2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgm2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgm2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgm2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgm2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgm2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	100.0	73	5	US-10-067-832D-2
2	394	100.0	73	5	US-10-067-832D-15
3	394	100.0	73	5	US-10-067-832D-16
4	394	100.0	73	5	US-10-965-898-68
5	394	100.0	115	4	US-10-264-049-2463
6	370	93.9	73	5	US-10-067-832D-19
7	366	92.9	68	4	US-10-424-599-250542
8	354	89.8	73	5	US-10-067-832D-20
9	339	86.0	73	5	US-11-097-143-6522
10	331	84.0	73	5	US-10-067-832D-17
11	319	81.0	73	4	US-10-424-599-212296
12	312	79.2	73	4	US-10-437-963-147593
13	312	79.2	73	4	US-10-767-701-46049
14	312	79.2	73	4	US-10-425-115-233843
15	312	79.2	73	4	US-10-425-115-233846
16	312	79.2	73	4	US-10-425-115-233848
17	312	79.2	73	4	US-10-425-115-233849
18	312	79.2	73	4	US-10-425-115-233851
19	312	79.2	73	4	US-10-425-115-233853
20	311	78.9	73	4	US-10-425-115-288480
21	311	78.9	73	4	US-10-425-115-288482
22	311	78.9	73	4	US-10-425-115-288483
23	311	78.9	73	4	US-10-425-115-288486
24	304.5	77.3	74	4	US-10-424-599-207467
25	300	76.1	73	4	US-10-425-115-288481
26	286	72.6	66	5	US-10-067-832D-18
27	285	72.3	73	5	US-10-739-930-8461

28	281	71.3	73	4	US-10-425-115-288479
29	252	64.0	89	4	US-10-083-357-770
30	252	64.0	320	4	US-10-338-411-11
31	252	64.0	320	4	US-10-389-640-11
32	248	62.9	73	4	US-10-425-115-290033
33	246	62.4	89	4	US-10-425-115-355452
34	242	61.4	82	4	US-10-424-599-191611
35	234	59.4	71	4	US-10-424-599-147197
36	194	49.2	110	4	US-10-425-115-302543
37	184	46.7	70	4	US-10-425-115-269824
38	171	43.4	79	4	US-10-425-115-269824
39	170	43.1	63	4	US-10-424-599-229016
40	163	41.4	33	5	US-10-067-832D-14
41	160.5	40.7	93	4	US-10-424-599-266457
42	149	37.8	45	4	US-10-425-115-262612
43	138	35.0	27	3	US-09-864-761-41766
44	138	35.0	55	4	US-10-425-115-270723
45	131	33.2	59	4	US-10-425-115-356308

ALIGNMENTS

RESULT 1

US-10-067-832D-2  
; Sequence 2, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEY  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent in Ver. 3.2  
; SEQ ID NO 2  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Psammomys obesus  
US-10-067-832D-2

Query Match 100.0%; Score 394; DB 5; Length 73;  
Best Local Similarity 100.0%; Pred. No. 1.5e-39;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MIEVVCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGRNKKIVLKKWYTFKDHVSLGDY 60
Db	1	MIEVVCNDRLGKVRKVCNTDDTIGDLKKLIAAQTGRNKKIVLKKWYTFKDHVSLGDY 60
Qy	61	EIHGDMNLELYQ 73
Db	61	EIHGDMNLELYQ 73

RESULT 2

US-10-067-832D-15  
; Sequence 15, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEY  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D

```
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 09/331,930
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP 0117
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP 0323
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 15
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-832D-15

Query Match          100.0%; Score 394; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 EIHGGMNLELYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYQ 73
   ||||||||||||||||

RESULT 3
US-10-067-832D-16
; Sequence 16, Application US/10067832D
; Publication No. US20050059108A1
; GENERAL INFORMATION:
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975200701
; CURRENT APPLICATION NUMBER: US/10/067,832D
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 09/331,930
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP 0117
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP 0323
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 16
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-067-832D-16

Query Match          100.0%; Score 394; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 EIHGGMNLELYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYQ 73
   ||||||||||||||||

RESULT 4
US-10-965-898-68
; Sequence 68, Application US/10965898
```

```
; Publication No. US20050084936A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Bandman, Olga
; Hillman, Jennifer L.
; Au-Young, Janice
; Tang, Y. Tom
; Yue, Henry
; Shan, Purvi
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
; NUMBER OF SEQUENCES: 150
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/965,898
; FILING DATE: 18-Oct-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,403
; FILING DATE: 31-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0455 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT19
; CLONE: 3245954
; SEQUENCE DESCRIPTION: SEQ ID NO: 68 :
US-10-965-898-68

Query Match          100.0%; Score 394; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MIEVNCNDRLGKVKVKNCTDDTIGDLKKLIAAQTGRNKNIVLKKWYTFKDHVSLGDY 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 EIHGGMNLELYQ 73
   ||||||||||||||||
Db 61 EIHGGMNLELYQ 73
   ||||||||||||||||

RESULT 5
US-10-264-049-2463
; Sequence 2463, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
```

; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 2463  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-2463

Query Match 100.0%; Score 394; DB 4; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.6e-39;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Db 43 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 102  
Qy 61 EIHGGMNLELYQ 73  
Db 103 EIHGGMNLELYQ 115

RESULT 6  
US-10-067-832D-19  
; Sequence 19, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975200701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 19  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-067-832D-19

Query Match 93.9%; Score 370; DB 5; Length 73;  
Best Local Similarity 93.1%; Pred. No. 1.2e-36;  
Matches 67; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Db 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Qy 61 EIHGGMNLELYY 72  
Db 61 EIHGGMNLELYY 72

RESULT 7  
US-10-424-599-250542  
; Sequence 250542, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 250542  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_6826C.1.pap  
US-10-424-599-250542

Query Match 92.9%; Score 366; DB 4; Length 68;  
Best Local Similarity 100.0%; Pred. No. 3.2e-36;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Db 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Qy 61 EIHGGMNLELYY 68  
Db 61 EIHGGMNLELYY 68

RESULT 8  
US-10-067-832D-20  
; Sequence 20, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975200701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 20  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-067-832D-20

Query Match 89.8%; Score 354; DB 5; Length 73;  
Best Local Similarity 90.3%; Pred. No. 9.8e-35;  
Matches 65; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Db 1 MIEVVCNDRLGKVRVKCNTDDTTIGDLKLLIAAQTGRWNKIVLKKWYTIKDHVSLG DY 60  
Qy 61 EIHGGMNLELYY 72  
Db 61 EIHGGMNLELYY 72

RESULT 9  
US-11-097-143-6522  
; Sequence 6522, Application US/11097143  
; Publication No. US2005020858A1

GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6522  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-6522

Query Match 86.0%; Score 339; DB 6; Length 73;  
Best Local Similarity 86.3%; Pred. No. 6.2e-33;  
Matches 63; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MIEVVCNDRLGKVRVKCNCTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTFKDHVSLG DY 60  
DB 1 MIEITCNDRLGKVRVKCNPDPTIGDLKKLIAAQTGTHKEKIVLKKWYTFKDPRLSDY 60  
QY 61 EIHGGMNLELYY 73  
DB 61 EIHGGMNLELYY 73

RESULT 10  
US-10-067-832D-17  
; Sequence 17, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975200701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 17  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-067-832D-17

Query Match 84.0%; Score 331; DB 5; Length 73;  
Best Local Similarity 80.8%; Pred. No. 5.7e-32;  
Matches 59; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MIEVVCNDRLGKVRVKCNCTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTFKDHVSLG DY 60  
DB 1 MIEITVNDRLGKVRVKCNPSDTIGDLKKLIAAQTGTRWEKIVLKKWYTIYKDHITLMDY 60  
QY 61 EIHGGMNLELYY 73  
DB 61 EIHGGMNLELYY 73  
RESULT 11  
US-10-424-599-212296  
; Sequence 212296, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei.  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 212296  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_33729C.1.pap  
US-10-424-599-212296

Query Match 81.0%; Score 319; DB 4; Length 73;  
Best Local Similarity 81.9%; Pred. No. 1.6e-30;  
Matches 59; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MIEVVCNDRLGKVRVKCNCTDDTIGDLKKLIAAQTGTRWNKIVLKKWYTFKDHVSLG DY 60  
DB 1 MIEVVCNDRLGKVRVKCNDDDTIGDLKKLIAAQTGTRADKIRIOKWYTIYKDHITLMDY 60  
QY 61 EIHGGMNLELYY 72  
DB 61 EIHGGMNLELYY 72  
RESULT 12  
US-10-437-963-147593  
; Sequence 147593, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 147593  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48107C.1.pep  
US-10-437-963-147593

Query Match 79.2%; Score 312; DB 4; Length 73;  
Best Local Similarity 80.6%; Pred. No. 1.1e-29;  
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKLLAAQTGTRWNKIVLKKWYTFKDHVSLGDY 60  
Db 1 MIEVVLNDRLGKVRKVCNEDDTIGDLKLLVAAQTGTRPEKIRIQKWNYYKHITLKY 60  
Qy 61 EIHGGMNLELY 72  
Db 61 EIHGGMLELY 72

RESULT 13  
US-10-767-701-46049  
; Sequence 46049, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 46049  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C594\_1.pep  
US-10-767-701-46049

Query Match 79.2%; Score 312; DB 4; Length 73;  
Best Local Similarity 80.6%; Pred. No. 1.1e-29;  
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKLLAAQTGTRWNKIVLKKWYTFKDHVSLGDY 60  
Db 1 MIEVVLNDRLGKVRKVCNEDDTIGDLKLLVAAQTGTRPEKIRIQKWNYYKHITLKY 60  
Qy 61 EIHGGMNLELY 72  
Db 61 EIHGGMLELY 72

RESULT 14  
US-10-425-115-233843  
; Sequence 233843, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 233843  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_144858C.1.pep  
US-10-425-115-233843

Query Match 79.2%; Score 312; DB 4; Length 73;  
Best Local Similarity 80.6%; Pred. No. 1.1e-29;  
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKLLAAQTGTRWNKIVLKKWYTFKDHVSLGDY 60  
Db 1 MIEVVLNDRLGKVRKVCNEDDTIGDLKLLVAAQTGTRPEKIRIQKWNYYKHITLKY 60  
Qy 61 EIHGGMNLELY 72  
Db 61 EIHGGMLELY 72

RESULT 15  
US-10-425-115-233846  
; Sequence 233846, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 233846  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_144860C.1.pep  
US-10-425-115-233846

Query Match 79.2%; Score 312; DB 4; Length 73;  
Best Local Similarity 80.6%; Pred. No. 1.1e-29;  
Matches 58; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 MIEVVCNDRLGKVRKVCNTDDTIGDLKLLAAQTGTRWNKIVLKKWYTFKDHVSLGDY 60  
Db 1 MIEVVLNDRLGKVRKVCNEDDTIGDLKLLVAAQTGTRPEKIRIQKWNYYKHITLKY 60  
Qy 61 EIHGGMNLELY 72  
Db 61 EIHGGMLELY 72

Search completed: April 8, 2006, 00:07:55  
Job time : 167 secs

**THIS PAGE BLANK (USPTO)**









```
Db          76 NSEFY 80

RESULT 10
US-11-096-568A-2440
; Sequence 2440, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2440
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(379)
; OTHER INFORMATION: Ceres Seq. ID no. 15221205
US-11-096-568A-2440

Query Match          14.2%; Score 56; DB 7; Length 379;
Best Local Similarity 30.8%; Pred. No. 25;
Matches 20; Conservative 9; Mismatches 18; Indels 18; Gaps 4;

QY 19 NTDTTIGDLKKLIAAQTGRN-----KIV-----LKKWYTFKDHVSLGDIETH---DGM 66
||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 24 NTDGVVGDILRKYLCKYKGQWDGRESLKIVGKTPLEAATVE-----DYELPCRVDVF 77
||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY 67 NLELY 71
|||
Db 78 NSEFY 82

RESULT 11
US-11-024-959-421
; Sequence 421, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 421
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-421

Query Match          14.2%; Score 56; DB 7; Length 731;
Best Local Similarity 26.2%; Pred. No. 54;
Matches 11; Conservative 10; Mismatches 21; Indels 21; Gaps 0;

QY 7 NDRLGKVRVKCNTDDTIGDLKKLIAAQTGRWNKIVLKKWY 48
||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 617 NDLLOENLNCGTGHEPLVDLSSIDAIRNGNSILELIKDSWY 658

RESULT 12
US-11-058-924-6
; Sequence 6, Application US/11058924
; Publication No. US20060019308A1
; GENERAL INFORMATION:
; APPLICANT: Lund, Frances E.
; APPLICANT: Randall, Troy D.
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: 13315/5
; CURRENT APPLICATION NUMBER: US/11/058,924
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 09/982,616
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 258
; TYPE: PRT
; ORGANISM: A. californica
US-11-058-924-6

Query Match          14.1%; Score 55.5; DB 7; Length 258;
Best Local Similarity 36.4%; Pred. No. 18;
Matches 12; Conservative 9; Mismatches 11; Indels 11; Gaps 1;

QY 2 IEVVCNDRLGKVRVKCNTDDTIGDLKKLIAAQ 34
::| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 190 VKVIVHLRLGKIEKCGAGSLL-DLEKLVKAK 221

RESULT 13
US-11-052-554A-222
; Sequence 222, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 222
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-222

Query Match          14.0%; Score 55; DB 7; Length 306;
Best Local Similarity 26.1%; Pred. No. 26;
Matches 24; Conservative 12; Mismatches 26; Indels 30; Gaps 4;

QY 1 MIEVVVCNDRLG-----KKVRVKCNTDDTIGDLKKLIAAQTGRWNKIVLKKWYT 49
::| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 LLLLVCLSLGLFACSTKKTADKLTVA-TNSIIADITKNIAG-----NKKVLHSIVP 58

QY 50 IFKD-----HVSGLGDIYEHGNNLE 69
::| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 59 VGRDPHEYPLEPDKVKTQADVIFVYNGINLE 90

RESULT 14
US-11-218-281-27
; Sequence 27, Application US/11218281
; Publication No. US20060024758A1
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:31:58 ; Search time 8.77373 Seconds  
(without alignments)  
1021.610 Million cell updates/sec

Title: US-10-067-832D-13  
Perfect score: 188  
Sequence: 1 atgatcaggtgtgttgcac.....agaagctgattgcagcctaa 102

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat\_07042006\_175609\_9682/app query.fasta.1  
-DB=A Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPC1=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USER=US10067832 @CGN 1.1 476 @runat\_07042006\_175609\_9682 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21:\*\*

- 1: Geneseqp1980s:\*\*
- 2: Geneseqp1990s:\*\*
- 3: Geneseqp2000s:\*\*
- 4: Geneseqp2001s:\*\*
- 5: Geneseqp2002s:\*\*
- 6: Geneseqp2003as:\*\*
- 7: Geneseqp2003bs:\*\*
- 8: Geneseqp2004s:\*\*
- 9: Geneseqp2005s:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	86.7	33	2 AAY08414	Human bea
2	163	86.7	33	8 ADR43346	Human Bea
3	163	86.7	73	3 AAY08413	Aay08413 P. obesus
4	163	86.7	73	3 AAB36290	Aab36290 Israeli s
5	163	86.7	73	4 AAM39680	Aam39680 Human pol
6	163	86.7	89	4 AAM41466	Aam41466 Human pol
7	163	86.7	115	5 ABP41331	Abp41331 Human ova
8	152	80.9	73	4 AAB59910	Aab59910 Drosophila
9	147.5	78.5	32	3 AAB36291	Aab36291 Human bea

10	146	77.7	73	3 AAG40885	Aag40885 Zea mays
11	146	77.7	73	3 AAG41925	Aag41925 Arabidops
12	146	77.7	96	3 AAG41924	Aag41924 Arabidops
13	141	75.0	73	3 AAG27281	Aag27281 Zea mays
14	130.5	69.4	73	8 ADT58384	Adt58384 plant pol
15	114	60.6	32	9 ADV95409	Adv95409 Yeast ubi
16	114	60.6	40	9 ADV95411	Adv95411 Yeast ubi
17	114	60.6	89	5 ADH32312	Adh32312 Novel yea
18	114	60.6	320	6 ABR83620	AbR83620 HUB1-GFP
19	86	45.7	104	5 ADH32742	Adh32742 Yeast smQ
20	83	44.1	110	8 ADY04784	Ady04784 plant ful
21	83	44.1	112	8 ADY06931	Ady06931 plant ful
22	83	44.1	112	8 ADX97089	Adx97089 plant ful
23	83	44.1	112	8 ADX78384	Adx78384 plant ful
24	83	44.1	113	8 ADY05953	Ady05953 plant ful
25	83	44.1	124	8 ADY04557	Ady04557 plant ful
26	83	44.1	128	8 ADX96955	Adx96955 plant ful
27	73	38.8	123	8 ADY05551	Ady05551 plant ful
28	73	38.8	125	8 ADY05562	Ady05562 plant ful
29	68	36.2	27	4 AAM21187	Aam21187 Peptide #
30	68	36.2	27	4 ABB43508	Abb43508 Peptide #
31	68	36.2	27	4 AAM37402	Aam37402 Peptide #
32	68	36.2	27	4 ABB26468	Abb26468 Protein #
33	68	36.2	27	4 AAM77251	Aam77251 Human bon
34	68	36.2	27	4 AAM64443	Aam64443 Human bra
35	68	36.2	27	5 ABG46267	Abg46267 Human pep
36	59	31.4	92	7 ABO80827	AbO80827 Pseudomon
37	58	30.9	1447	3 AAB36514	Aab36514 Candida a
38	58	30.9	1447	5 ABP73835	Abp73835 Candida a
39	56.5	30.1	282	2 AAR34706	Aar34706 NAD cycla
40	56.5	30.1	282	2 AAR70095	Aar70095 NAD cycla
41	56.5	30.1	282	2 AAM80290	Aam80290 Aplysia c
42	55.5	29.5	261	9 ADU91841	Adu91841 Carbonic
43	55.5	29.5	2000	8 ADP24787	Adp24787 PRO polyP
44	55.5	29.5	2000	9 ADY15157	Ady15157 PRO polyP
45	55.5	29.5	2695	2 AAM07632	Aam07632 Human cyp

ALIGNMENTS

RESULT 1

AAY08414  
ID AAY08414 standard; protein; 33 AA.

XX AC AAY08414;

XX AC

DT 24-JUL-1999 (first entry)

XX

DE Human beacon protein.

XX

XX Beacon; hypothalamus; obese; lean; agonist; antagonist; treatment;  
XX obesity; anorexia; weight maintenance; energy imbalance; diabetes;  
XX metabolic syndrome; dyslipidemia; hypertension; insulin resistance;  
XX medication; livestock; diagnosis; human.

XX Homo sapiens.

XX OS

XX Key Location/Qualifiers

FT Misc-difference 15

FT /label= unknown

FT /note= "encoded by CNC"

XX WO9923217-A1.

XX

PD 14-MAY-1999.

XX

XX 30-OCT-1998; 98WO-AU000902.

PF

PR 31-OCT-1997; 97AU-00000117.

PR 11-NOV-1997; 97AU-00000323.

XX (ITDI-) INT DIABETES INST.

PA (UYDE-) UNIV DEAKIN.

XX Zimmet PZ, Collier G;  
 XX WPI; 1999-337484/28.  
 DR N-PSDB; AAX57370.  
 XX New gene encoding a beacon protein associated with modulation of obesity,  
 PT diabetes and metabolic energy levels.  
 XX Claim 2; Page 54; 85pp; English.  
 XX This invention describes a novel beacon protein and its encoding nucleic  
 CC acid which is expressed in larger amounts in hypothalamus tissue of obese  
 CC animals compared to lean animals. Agonists and antagonists of beacon can  
 CC be used to treat obesity, anorexia, weight maintenance, energy imbalance,  
 CC diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin  
 CC resistance. The beacon protein, itself is used to manufacture medicaments  
 CC for treatment of obesity, anorexia, energy imbalance or diabetes. The  
 CC treatment is contemplated for both human and animals, such as those  
 CC important to the livestock industry. The antibody and polynucleotides are  
 CC useful in diagnosis of conditions as above  
 XX SQ Sequence 33 AA;

Alignment Scores:  
 Pred. No.: 4.39e-15 Length: 33  
 Score: 163.00 Matches: 33  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 86.7% Indels: 0  
 DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x AAY08414 (1-33)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAGTCCNCGTTAAATGCAACG 60  
 |||||  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysVal\*\*VallLysCysAsnThr 20  
 QY 61 GATGATACCATCGGGACCTTAAGAGCTGATTGCAGCC 99  
 |||||  
 Db 21 AspAspThrileGlyAspLeuLysLysLeuileAlaala 33

RESULT 2  
 ID ADR43346  
 AC ADR43346 standard; peptide; 33 AA.

XX ADR43346;  
 XX 04-NOV-2004 (first entry)  
 XX Human Beacon protein.  
 XX Beacon; Anorectic; Anabolic; Antidiabetic; Cytostatic; Cardiant;  
 KW Antiinflammatory; Immunosuppressive; Antiinfertility; Neuroprotective;  
 KW Nootropic; CLK; myopathy; obesity; anorexia; diabetes; cancer;  
 KW heart disease; inflammation; Alzheimer's; Parkinson's; Pearson Syndrome;  
 KW Pyruvate Carboxylase Deficiency; Pyruvate Dehydrogenase Deficiency;  
 KW Cancer; Acoustic Neuroma; Acute Lymphocytic Leukemia; Carcinoma;  
 KW Leukemia.  
 XX Homo sapiens.  
 XX W02004069866-A1.  
 XX 19-AUG-2004.  
 XX 10-FEB-2004; 2004WO-AU000147.  
 XX 10-FEB-2003; 2003US-0446191P.  
 XX (AUTO-) AUTOGEN RES PTY LTD.  
 XX (UYDE-) UNIV DEAKIN.

PI Collier G, Walder K, Kerr-Bayles L;  
 XX WPI; 2004-604412/58.  
 XX New isolated ligands of mammalian or avian Beacon, useful for e.g.  
 PT preventing or treating disorders associated with myopathy, obesity,  
 PT diabetes, cancer, heart disease, inflammation, or disorders associated  
 PT with the immune system.  
 XX Claim 3; SEQ ID NO 4; 194pp; English.

XX The present invention relates to an isolated ligand of mammalian or avian  
 CC Beacon or a homolog or derivative of the Beacon. The composition  
 CC (including the agent capable of modulating the interaction between a  
 CC Beacon and a CLK) or methods are useful for manufacturing a medicament  
 CC for the treatment of a condition characterized by a healthy or unhealthy  
 CC state, including the presence or absence of a disorder associated with  
 CC myopathy, obesity, anorexia, weight maintenance, diabetes, disorders  
 CC associated with mitochondrial dysfunction, genetic disorders, cancer,  
 CC heart disease, inflammation, disorders associated with the immune system,  
 CC infertility, disease associated with the brain and/or metabolic energy  
 CC levels. The disease is selected from Alzheimer's, Parkinson's, diabetes,  
 CC autism, and the aging process, LIC (Lethal Infantile Cardiomyopathy),  
 CC Abgr/-oxidation Defects, COX Deficiency, Mitochondrial Cytopathy, Alper's  
 CC Disease, Barth syndrome, Carnitine-Acyl-Carnitine Deficiency, Carnitine  
 CC Deficiency, Co-Enzyme Q10 Deficiency, Complex I Deficiency, Complex II  
 CC Deficiency, Complex III Deficiency, Complex IV Deficiency, Complex V  
 CC Deficiency, chronic progressive external ophthalmoplegia syndrome (CPEO),  
 CC CPT I Deficiency, Glutaric Aciduria Type II, Kearns-Sayre syndrome (KSS),  
 CC lactic acidosis, long-chain acyl-CoA dehydrogenase deficiency (LCAD),  
 CC LCHAD, Leigh Disease, Leber Hereditary Optic Neuropathy (LHON), Luft  
 CC Disease, mitochondrial DNA depletion, Mitochondrial Encephalopathy,  
 CC Pearson Syndrome, Pyruvate Carboxylase Deficiency, Pyruvate Dehydrogenase  
 CC Deficiency, and the other diseases mentioned in the specification. The  
 CC cancer is selected from ABL1 protooncogene, AIDS Related Cancers,  
 CC Acoustic Neuroma, Acute Lymphocytic Leukemia, Acute Myeloid Leukemia,  
 CC Adenocystic carcinoma, Adrenocortical Cancer, Agnogenic myeloid  
 CC metaplasia, Alopecia, Alveolar soft-part sarcoma, Anal cancer,  
 CC Angiosarcoma, Aplastic Anemia, Astrocytoma, Ataxia-telangiectasia, Basal  
 CC Cell Carcinoma (Skin), Bladder Cancer, Bone Cancer, Bowel cancer, Brain  
 CC Stem Glioma, Brain and CNS Tumours, Breast Cancer, CNS tumour, Carcinoid  
 CC Tumours, Cervical Cancer, Childhood Brain Tumour, Childhood Cancer,  
 CC Childhood Leukemia, and other cancers mentioned in the specification. The  
 CC 11 sequences mentioned in the body of the specification do not correspond  
 CC to the sequences represented in the SEQ ID listing.

XX SQ Sequence 33 AA;

Alignment Scores:  
 Pred. No.: 4.39e-15 Length: 33  
 Score: 163.00 Matches: 33  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 86.7% Indels: 0  
 DB: 8 Gaps: 0

US-10-067-832D-13 (1-102) x ADR43346 (1-33)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAGTCCNCGTTAAATGCAACG 60  
 |||||  
 Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysVal\*\*VallLysCysAsnThr 20  
 QY 61 GATGATACCATCGGGACCTTAAGAGCTGATTGCAGCC 99  
 |||||  
 Db 21 AspAspThrileGlyAspLeuLysLysLeuileAlaala 33

RESULT 3  
 AAY08413  
 ID AAY08413 standard; protein; 73 AA.  
 XX AAY08413;  
 XX 24-JUL-1999 (first entry)



XX DE P. obesus beacon protein.  
XX KW Beacon; hypothalamus; obese; lean; agonist; antagonist; treatment;  
XX KW obesity; anorexia; weight maintenance; energy imbalance; diabetes;  
KW metabolic syndrome; dyslipidemia; hypertension; insulin resistance;  
KW medicament; livestock; diagnosis.  
XX OS Psammomys obesus.  
XX KW WO9923217-A1.  
XX PN 14-MAY-1999.  
XX PD 30-OCT-1998; 98WO-AU000902.  
XX PF 31-OCT-1997; 97AU-00000117.  
XX PR 11-NOV-1997; 97AU-00000323.  
XX PR (ITDI-) INT DIABETES INST.  
XX PA (UYDE-) UNIV DEAKIN.  
XX PI Zimmet PZ, Collier G;  
XX PI WPI; 1999-337484/28.  
XX DR N-PSDB; AAX57359.  
XX DR New gene encoding a beacon protein associated with modulation of obesity,  
XX PT diabetes and metabolic energy levels.  
XX PS Claim 2; Page 50; 85pp; English.  
XX CC This invention describes a novel beacon protein and its encoding nucleic  
CC acid which is expressed in larger amounts in hypothalamus tissue of obese  
CC animals compared to lean animals. Agonists and antagonists of beacon can  
CC be used to treat obesity, anorexia, weight maintenance, energy imbalance,  
CC diabetes, metabolic syndrome, dyslipidemia, hypertension and/or insulin  
CC resistance. The beacon protein, itself is used to manufacture medicaments  
CC for treatment of obesity, anorexia, energy imbalance or diabetes. The  
CC treatment is contemplated for both human and animals, such as those  
CC important to the livestock industry. The antibody and polynucleotides are  
CC useful in diagnosis of conditions as above  
XX SQ Sequence 73 AA;  
XX  
Alignment Scores:  
Pred. No.: 4.74e-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 2 Gaps: 0  
US-10-067-832D-13 (1-102) x AAY08413 (1-73)  
Qy 1 ATGATCGAGGTGTTTGCACACCGTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
RESULT 4  
AAB36290  
ID AAB36290 standard; protein; 73 AA.  
XX AAB36290;  
XX 23-FEB-2001 (first entry)  
DT 23-FEB-2001 (first entry)  
DE Israeli sand rat beacon ligand.  
XX

KW Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;  
KW ligand.  
XX OS Psammomys obesus.  
XX PN WO200064931-A1.  
XX PD 02-NOV-2000.  
XX PF 19-APR-2000; 2000WO-AU000342.  
XX PR 23-APR-1999; 99AU-00009919.  
XX PR 24-MAR-2000; 2000AU-00006454.  
XX PA (AUTO-) AUTOGEN PTY LTD.  
XX PI Collier G, Walder K, Zimmet P;  
XX PI WPI; 2000-687311/67.  
XX DR N-PSDB; AAC81767.  
XX PT Ligand of beacon protein useful for treating obesity, anorexia, energy  
XX PT imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and  
XX PT insulin resistance.  
XX PS Claim 3; Fig 1; 67pp; English.  
XX CC The present invention is related to the isolation of a ligand known as  
XX CC beacon from the Israeli sand rat. Beacon is associated with the  
XX CC regulation of energy balance, and the protein, its coding sequence and  
XX CC analogues can be used in the treatment of diabetes, obesity, anorexia,  
XX CC energy imbalance, metabolic syndrome, dyslipidemia, hypertension and  
XX CC insulin resistance. In addition, they can be used in agriculture to  
XX CC produce leaner animals  
XX SQ Sequence 73 AA;  
XX  
Alignment Scores:  
Pred. No.: 4.74e-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 3 Gaps: 0  
US-10-067-832D-13 (1-102) x AAB36290 (1-73)  
Qy 1 ATGATCGAGGTGTTTGCACACCGTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
RESULT 5  
AAB39680  
ID AAB39680 standard; protein; 73 AA.  
XX AAB39680;  
XX 22-OCT-2001 (first entry)  
DT 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 2825.  
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX Homo sapiens.  
XX

XX WO200153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US034263.  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 23-NOV-2000; 2000US-00727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI58836.  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX Example 4; SEQ ID NO 2825; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AM38642-AM42213) with nootropic,  
XX immunosuppressant and cytotostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX SQ Sequence 73 AA;  
Alignment Scores:  
Pred. No.: 4.74e-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 4 Gaps: 0  
US-10-067-832D-13 (1-102) x AAM39680 (1-73)  
QY 1 ATGATCGAGGTGTTGTTGCAACGACCGCTCTGGGGAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetIleGluValValCysAsnApgArgLeuGlyLysValargVallysCysAsnThr 20  
QY 61 GATGATACCATCGGGGACCTTAAGACCTGATTGCAGCC 99  
Db 21 AspApthrIleGlyApgLeuLysLysLeuIleAlaAala 33  
RESULT 6  
AAM41466  
ID AAM41466 standard; protein; 89 AA.  
XX  
XX AAM41466;  
XX

DT 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 6397.  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX leukaemia.  
XX Homo sapiens.  
XX WO200153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US034263.  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 23-NOV-2000; 2000US-00727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI60622.  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX Example 2; SEQ ID NO 6397; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AM38642-AM42213) with nootropic,  
XX immunosuppressant and cytotostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX SQ Sequence 89 AA;  
Alignment Scores:  
Pred. No.: 4.83e-15 Length: 89  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 4 Gaps: 0  
US-10-067-832D-13 (1-102) x AAM41466 (1-89)  
QY 1 ATGATCGAGGTGTTGTTGCAACGACCGCTCTGGGGAAGGTCNCGTTAAATGCAACACG 60  
|||||

Db 17 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 36

Qy 61 GATGATACCATCGGGGACCTTAAGAGCTGATTGCAGCC 99  
|||||

Db 37 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 49  
|||||

RESULT 7

ABP41331

ID ABP41331 standard; protein; 115 AA.

XX ABP41331;

AC

XX

XX

DT 22-AUG-2002 (first entry)

XX

DE Human ovarian antigen HACNC39, SEQ ID NO:2463.

XX

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KW antiinflammatory; gynaecological; reproductive.

XX

OS Homo sapiens.

XX

XX

XX WO200200677-A1.

XX

XX

PD 03-JAN-2002.

XX

XX

XX 07-JUN-2001; 2001WO-US018569.

XX

XX

XX 07-JUN-2000; 2000US-0209467P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Birse CE, Rosen CA;

XX

XX WPI; 2002-147878/19.

XX

XX N-PSDB; ABQ54408.

XX

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian

PT cancer), immune disorders, cardiovascular disorders and neurological

PT diseases.

XX

PS Claim 11; SEQ ID NO 2463; 2922pp; English.

XX

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-

CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

CC treating, prognosing or preventing various ovary and/or breast-related

CC disorders. Such conditions include ovarian cancer and breast cancer, and

CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

CC respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and

CC polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the

CC

CC identification of individuals and in forensic analysis, and the

CC polypeptides may be used as food additives or to prepare antibodies

CC useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 115 AA;

Alignment Scores:

Pred. No.: 4.95e-15 Length: 115

Score: 163.00 Matches: 32

Percent Similarity: 97.0% Conservative: 0

Best Local Similarity: 97.0% Mismatches: 1

Query Match: 86.7% Indels: 0

DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x ABP41331 (1-115)

Qy 1 ATGATCGAGGTGTTTTCACACGCCGCTCGGGGAAAAAGTCCNGTTAAATGCACACG 60  
|||||

Db 43 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 62  
|||||

Qy 61 GATGATACCATCGGGGACCTTAAGAGCTGATTGCAGCC 99  
|||||

Db 63 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 75  
|||||

RESULT 8

ABBS59910

ID ABBS59910 standard; protein; 73 AA.

XX

XX

AC ABBS59910;

XX

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 6522.

XX

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX

XX Drosophila melanogaster.

XX

OS

XX WO200171042-A2.

XX

XX

PD 27-SEP-2001.

XX

XX

XX 23-MAR-2001; 2001WO-US009231.

XX

XX

XX 23-MAR-2000; 2000US-0191637P.

PR

PR 11-JUL-2000; 2000US-00614150.

XX

XX (PEKE ) PE CORP NY.

XX

XX Venter JC, Adams M, Li PWD, Myers EW;

PI

XX

XX WPI; 2001-656860/75.

DR

DR N-PSDB; ABL04013.

XX

XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

XX

PS Disclosure; SEQ ID NO 6522; 21pp + Sequence Listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-

CC ABBS72072). The sequence data for this patent did not form part of the

CC

CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 73 AA;

Alignment Scores:  
Pred. No.: 2.01e-13 Length: 73  
Score: 152.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 80.9% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-13 (1-102) x ABB59910 (1-73)

QY 1 ATGATCGAGGTGTTTTCGAACGACCGTCTGGGAAAAGCTCCNCGTTAAATGCAACAG 60  
Dd |||||:|||||  
1 MetIleGluIleThrCysAsnAspArgLeuGlyLysValArgVallysCysAsnPro 20  
QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99  
Dd |||||  
21 AspAspThrIleGlyAspLeuLysLeuIleAlaAla 33

RESULT 9  
AAB36291  
ID AAB36291 standard; protein; 32 AA.  
XX  
AC AAB36291;  
XX  
DT 23-FEB-2001 (first entry)  
XX  
DE Human beacon (short form) protein sequence.  
XX  
KW Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;  
KW ligand; human.  
XX  
OS Homo sapiens.  
XX  
PN WQ200064931-A1.  
XX  
PD 02-NOV-2000.  
XX  
PF 19-APR-2000; 2000WO-AU000342.  
XX  
PR 23-APR-1999; 99AU-00009919.  
PR 24-MAR-2000; 2000AU-00006454.  
XX  
PA (AUTO-) AUTOGEN PTY LTD.  
XX  
PI Collier G, Walder K, Zimmet P;  
XX  
DR WPI; 2000-687311/67.  
DR N-PSDB; AAC81770.  
XX  
XX Ligand of beacon protein useful for treating obesity, anorexia, energy  
PT imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and  
PT insulin resistance.  
XX  
PS Disclosure; Fig 1B; 67pp; English.  
XX  
CC The present invention is related to the isolation of a ligand known as  
CC beacon from the Israeli sand rat. Beacon is associated with the  
CC regulation of energy balance, and the protein, its coding sequence and  
CC analogues can be used in the treatment of diabetes, obesity, anorexia,  
CC energy imbalance, metabolic syndrome, dyslipidaemia, hypertension and  
CC insulin resistance. In addition, they can be used in agriculture to  
CC produce leaner animals  
XX  
SQ Sequence 32 AA;

Alignment Scores:  
Pred. No.: 8.61e-13 Length: 32  
Score: 147.50 Matches: 32

Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 0  
Query Match: 78.5% Indels: 1  
DB: 3 Gaps: 1

US-10-067-832D-13 (1-102) x AAB36291 (1-32)

QY 1 ATGATCGAGGTGTTTTCGAACGACCGTCTGGGAAAAGCTCCNCGTTAAATGCAACAG 60  
Dd |||||  
1 MetIleGluValValCysAsnAspArgLeuGlyLysVal\*\*\*VallysCysAsnThr 20  
QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99  
Dd |||||  
21 AspAspThrIleGlyAspLeu---LysLeuIleAlaAla 32

RESULT 10  
AAG40885  
ID AAG40885 standard; protein; 73 AA.  
XX  
AC AAG40885;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Zea mays protein fragment SEQ ID NO: 50790.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143452P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148665P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 16-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 28-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

```
Pred. No.: 1 55e-12 Length: 73
Score: 146.00 Matches: 29
Percent Similarity: 90.9% Conservative: 1
Best Local Similarity: 87.9% Mismatches: 3
Query Match: 77.7% Indels: 0
DB: 3 Gaps: 0

US-10-067-832D-13 (1-102) x AAG40885 (1-73)

QY 1 ATGATCGAGTTGTTTGCACACGCGTCTGGGGAAGGTCNCGTTAAATGCAACACG 60
    |||||
Db 1 MetIleGluValValLeuAsnAspArgLeuGlyLysValArgValLysCysAsnGlu 20
    |||||

QY 61 GATGATACCATCGGACCTTAAGACCTGATTGCACCC 99
    |||||
Db 21 AspAspThrIleGlyAspLeuLysLysLeuValAlaAla 33

RESULT 11
AAG41925
ID AAG41925 standard; protein; 73 AA.
XX
AC AAG41925;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52222.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136031P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136393P.
PR 05-JUN-1999; 99US-0137222P.
PR 09-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
```

```
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 07-AUG-1999; 99US-0147303P.
PR 08-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 14-AUG-1999; 99US-0148565P.
PR 15-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 19-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 21-AUG-1999; 99US-0149929P.
PR 22-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 24-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151065P.
PR 28-AUG-1999; 99US-0151080P.
PR 29-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 02-SEP-1999; 99US-0153070P.
PR 03-SEP-1999; 99US-0153758P.
PR 04-SEP-1999; 99US-0154018P.
PR 05-SEP-1999; 99US-0154039P.
PR 06-SEP-1999; 99US-0154779P.
PR 07-SEP-1999; 99US-0155139P.
PR 08-SEP-1999; 99US-0155486P.
PR 09-SEP-1999; 99US-0155659P.
PR 10-SEP-1999; 99US-0156458P.
PR 11-SEP-1999; 99US-0157117P.
PR 12-SEP-1999; 99US-0157753P.
PR 13-SEP-1999; 99US-0157865P.
PR 14-SEP-1999; 99US-0158029P.
PR 15-SEP-1999; 99US-0158232P.
PR 16-SEP-1999; 99US-0158369P.
PR 17-SEP-1999; 99US-0159293P.
PR 18-SEP-1999; 99US-0159294P.
PR 19-SEP-1999; 99US-0159295P.
PR 20-SEP-1999; 99US-0159329P.
PR 21-SEP-1999; 99US-0159330P.
PR 22-SEP-1999; 99US-0159331P.
PR 23-SEP-1999; 99US-0159637P.
PR 24-SEP-1999; 99US-0159638P.
PR 25-SEP-1999; 99US-0159584P.
PR 26-SEP-1999; 99US-0160741P.
PR 27-SEP-1999; 99US-0160767P.
PR 28-SEP-1999; 99US-0160768P.
PR 29-SEP-1999; 99US-0160770P.
PR 30-SEP-1999; 99US-0160814P.
PR 01-OCT-1999; 99US-0160815P.
PR 02-OCT-1999; 99US-0160980P.
PR 03-OCT-1999; 99US-0160981P.
PR 04-OCT-1999; 99US-0160989P.
PR 05-OCT-1999; 99US-0161404P.
PR 06-OCT-1999; 99US-0161405P.
PR 07-OCT-1999; 99US-0161406P.
PR 08-OCT-1999; 99US-0161359P.
PR 09-OCT-1999; 99US-0161360P.
PR 10-OCT-1999; 99US-0161361P.
PR 11-OCT-1999; 99US-0161920P.
PR 12-OCT-1999; 99US-0161921P.
PR 13-OCT-1999; 99US-0161922P.
PR 14-OCT-1999; 99US-0161923P.
PR 15-OCT-1999; 99US-0162142P.

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 07-AUG-1999; 99US-0147303P.
PR 08-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 14-AUG-1999; 99US-0148565P.
PR 15-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 19-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 21-AUG-1999; 99US-0149929P.
PR 22-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 24-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151065P.
PR 28-AUG-1999; 99US-0151080P.
PR 29-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 02-SEP-1999; 99US-0153070P.
PR 03-SEP-1999; 99US-0153758P.
PR 04-SEP-1999; 99US-0154018P.
PR 05-SEP-1999; 99US-0154039P.
PR 06-SEP-1999; 99US-0154779P.
PR 07-SEP-1999; 99US-0155139P.
PR 08-SEP-1999; 99US-0155486P.
PR 09-SEP-1999; 99US-0155659P.
PR 10-SEP-1999; 99US-0156458P.
PR 11-SEP-1999; 99US-0157117P.
PR 12-SEP-1999; 99US-0157753P.
PR 13-SEP-1999; 99US-0157865P.
PR 14-SEP-1999; 99US-0158029P.
PR 15-SEP-1999; 99US-0158232P.
PR 16-SEP-1999; 99US-0158369P.
PR 17-SEP-1999; 99US-0159293P.
PR 18-SEP-1999; 99US-0159294P.
PR 19-SEP-1999; 99US-0159295P.
PR 20-SEP-1999; 99US-0159329P.
PR 21-SEP-1999; 99US-0159330P.
PR 22-SEP-1999; 99US-0159331P.
PR 23-SEP-1999; 99US-0159637P.
PR 24-SEP-1999; 99US-0159638P.
PR 25-SEP-1999; 99US-0159584P.
PR 26-SEP-1999; 99US-0160741P.
PR 27-SEP-1999; 99US-0160767P.
PR 28-SEP-1999; 99US-0160768P.
PR 29-SEP-1999; 99US-0160770P.
PR 30-SEP-1999; 99US-0160814P.
PR 01-OCT-1999; 99US-0160815P.
PR 02-OCT-1999; 99US-0160980P.
PR 03-OCT-1999; 99US-0160981P.
PR 04-OCT-1999; 99US-0160989P.
PR 05-OCT-1999; 99US-0161404P.
PR 06-OCT-1999; 99US-0161405P.
PR 07-OCT-1999; 99US-0161406P.
PR 08-OCT-1999; 99US-0161359P.
PR 09-OCT-1999; 99US-0161360P.
PR 10-OCT-1999; 99US-0161361P.
PR 11-OCT-1999; 99US-0161920P.
PR 12-OCT-1999; 99US-0161921P.
PR 13-OCT-1999; 99US-0161922P.
PR 14-OCT-1999; 99US-0161923P.
PR 15-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.55e-12 Length: 73
Score: 146.00 Matches: 29
Percent Similarity: 90.9% Conservative: 1
Best Local Similarity: 87.9% Mismatches: 3
Query Match: 77.7% Indels: 0
DB: 3 Gaps: 0

US-10-067-832D-13 (1-102) x AAG41925 (1-73)

Qy 1 ATGATCGAGTTGTTTCACAGCGCTCGGGGAAAAGCTCCNGCTTAATGCAACACG 60
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysLysValArgValLysCysAenAap 20
Qy 61 GATGATACCATCGGGGACCTTAAAGAGCTGATTTCAGGCC 99
Db 21 AspAspThrileGlyAspLeuLysLysLeuValAlaAla 33

RESULT 12
AAG41924
ID AAG41924 standard; protein; 96 AA.
AC AAG41924;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52221.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
```

PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140981P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141824P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142052P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145112P.  
PR 22-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 15-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.



PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores: 1.6e-12 Length: 96  
Pred. No.: 146.00 Matches: 29  
Score: 90.9% Conservaive: 1  
Percent Similarity: 87.9% Mismatches: 3  
Best Local Similarity: 77.7% Indels: 0  
Query Match: 3 Gaps: 0  
DB:

US-10-067-832D-13 (1-102) x AAG41924 (1-96)

QY 1 ATGATCGAGTGTGTTTGCACGACCGTCTGGGAAAAGTCCNCGTTAATGCAACAGC 60  
|||||  
Db 24 MetileGluValValLeuAsnAspArgLeuGlyLysLysValArgValLysCysAsnAsp 43  
|||||  
QY 61 GATGATACCATCGGGGACCTTAAAGAAGCTGATTCAGCC 99  
|||||  
Db 44 AspAspThrileGlyAspLeuLysLysLeuValAlaAla 56  
|||||

RESULT 13

AAG27281

ID AAG27281 standard; protein; 73 AA.

AC AAG27281;

XX DT 17-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 32055.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.

XX Zea mays subep. mays.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 23-MAR-1999; 99US-0126264P.

XX 23-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140891P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.

```
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 8 54e-12 Length: 73
Score: 141.00 Matches: 28
Percent Similarity: 87.9% Conservative: 1
Best Local Similarity: 84.8% Mismatches: 4
Query Match: 75.0% Indels: 0
DB: 3 Gaps: 0

US-10-067-832D-13 (1-102) x AAG27281 (1-73)

QY 1 ATGATCGAGGTGTTTGGCAACGACCTCTGGGGAAAAGTCNCCTTAATCAACACG 60
Db 1 MetIleGluValValLeuAsnAspArgLeuGlyLysLysValArgValLysCysAsnGlu 20
QY 61 GATGATACCATCGGGGACCTTAAGAAAGCTGATTGCAGCC 99
Db 21 AspAspThrIleGlyAspLeuAsnLysLeuValAlaAla 33

RESULT 14
ADT58384
ID ADT58384 standard; protein; 73 AA.
XX AC ADT58384;
XX DT 13-JAN-2005 (first entry)
XX DE Plant polypeptide, SEQ ID 8461.
XX KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
disease resistance; galactomannan production; plant growth regulator;
heat tolerance; herbicide tolerance; lignin production;
extreme osmotic condition tolerance; pathogens resistance;
pest resistance; yield improvement; seed oil yield; seed protein yield.
XX OS Viridiplantae.
XX PN US2004216190-A1.
XX PD 28-OCT-2004.
XX PF 18-DEC-2003; 2003US-00739930.
XX PR 28-APR-2003; 2003US-00424599.
XX PR 28-APR-2003; 2003US-00425115.
XX PA (KOVA/) KOVALIC D K.
XX PI Kovalic DK;
XX DR WPI; 2004-757369/74.
XX PT New recombinant DNA constructs useful in the field of biochemistry and
genetics, and in particular for producing transgenic plants with improved
biological characteristics.
XX PS Claim 2; SEQ ID NO 8461; 14pp; English.
XX CC The invention relates a recombinant DNA construct comprising a
polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
(SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
Arabidopsis, wheat and rape but the specification does not indicate which
sequences is derived from which organism. Also included is a method of
producing a plant having an improved property, comprising transforming a
plant with a recombinant DNA construct comprising a promoter region
functional in a plant cell operably joined to a polynucleotide encoding a
polypeptide associated with the property, and growing the transformed
plant. The property is selected from improving plant cold tolerance, for
```

CC manipulating growth rate in plant cells by modification of the cell cycle  
 CC pathway, for improving plant drought tolerance, for providing increased  
 CC resistance to plant disease, for galactomannan production, for production  
 CC of plant growth regulators, for improving plant heat tolerance, for  
 CC improving plant tolerance to herbicides, for increasing the rate of  
 CC homologous recombination in plants, for lignin production, for improving  
 CC plant tolerance to extreme osmotic conditions, for improving plant  
 CC tolerance to pathogens or pests, for yield improvement by modification of  
 CC photosynthesis, for modifying seed oil yield and/or content, for  
 CC modifying seed protein yield and/or content, for yield improvement by  
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC and for yield improvement by providing improved plant growth and  
 CC development under at least one stress condition. The polynucleotide may  
 CC also encode a plant transcription factor. The methods and compositions of  
 CC the present invention are useful in the field of biochemistry and  
 CC genetics, in particular for producing transgenic plants with improved  
 CC biological characteristics such as increased yield, improved nitrogen  
 CC flow, increasing plant tolerance to cold or heat, improving plant  
 CC tolerance to extreme osmotic and drought conditions, and improving plant  
 CC tolerance to plant pests or pathogens. They can also be used in physical  
 CC arrays of molecules, plant breeding markers, computer-based storage and  
 CC analysis systems. The present sequence is one of the 5544 plant protein  
 CC sequences of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 73 AA;

Alignment Scores:  
 Pred. No.: 3,06e-10 Length: 73  
 Score: 130.50 Matches: 28  
 Percent Similarity: 87.9% Conservative: 1  
 Best Local Similarity: 84.8% Mismatches: 3  
 Query Match: 69.4% Indels: 1  
 DB: 8 Gaps: 1

US-10-067-832D-13 (1-102) x ADF58384 (1-73)

QY 1 ATGATCAGGTGTTTTCACACGACCGTCTGGGAAAAGTCCNCGTTAATGCAACAG 60  
 DB 1 MetilegluValValLeuAsnAspArgLeuGlyLysLysValArgValLysCys----\*\* 19  
 QY 61 GATGATACCATCGGGACCTTAGAAGCTGATTCGACC 99  
 DB 20 AspAspThrileGlyAspLeuLysLysLysValAlaAla 32

RESULT 15

ADV95409  
 ID ADV95409 standard; peptide; 32 AA.

AC ADV95409;

XX 07-APR-2005 (first entry)

DE Yeast ubiquitin like protein Hub1 amino terminal domain 1.

XX Fusion protein; protein folding; protein structure;  
 XX ubiquitin like protein.

XX Saccharomyces cerevisiae.

XX WO2005003313-A2.

XX 13-JAN-2005.

XX 28-JUN-2004; 2004WO-US020778.

XX 26-JUN-2003; 2003US-0482817P.

XX (LIFE-) LIFESENSORS INC.

XX Butt TR, Malakhov MP, Malakhova OA;

XX  
 DR  
 XX

WPI; 2005-091798/10.

Enhancing expression of target protein in host cell, by expressing  
 nucleic acid construct containing sequences encoding carboxy-terminal  
 domain of ubiquitin-like molecule and target protein and encoding fusion  
 protein in host cell.

Example 7; SEQ ID NO 41; 107pp; English.

The invention relates to enhancing (M1) the expression of a protein of  
 interest in a host cell, involving expressing a nucleic acid construct  
 comprising a nucleic acid sequence encoding a carboxy-terminal domain of  
 a ubiquitin-like (Ubl) molecule and a protein of interest, where the  
 presence of the carboxy-terminal domain of the Ubl molecule in the fusion  
 protein increases expression level of protein of interest in the host  
 cell. Also included are producing (M2) a protein of interest (involves  
 carrying out (M1), where the fusion protein, when expressed, comprises  
 the carboxy-terminal domain of Ubl molecule attached to the amino-  
 terminus of the protein of interest such that a cleavage site of the Ubl  
 molecule is immediately amino terminal to the protein of interest,  
 expressing the nucleic acid construct in the host cell, contacting the  
 expressed fusion protein with an amino-terminal domain of the same Ubl  
 molecule, thus generating a reconstituted Ubl, and contacting the  
 reconstituted Ubl with a protease specific to the Ubl molecule, thus  
 cleaving the fusion protein such that the protein of interest is now  
 produced), purifying (M3) the protein of interest (involving carrying out  
 specifically binding a purification tag (e.g. 6xHis tag or S-tag), and  
 contacting the purified reconstituted Ubl with a protease specific to the  
 Ubl molecule and cleaving the fusion protein such that the protein of  
 interest is now purified, or providing a nucleic acid construct which  
 encodes a fusion protein, containing a nucleic acid sequence encoding a  
 carboxy-terminal domain of Ubl molecule, the protein of interest, and a  
 first purification tag (attached to the amino-terminus of the carboxy-  
 terminal domain of a Ubl molecule) such that cleavage site of the Ubl  
 molecule is immediately amino terminal to the protein of interest,  
 expressing the nucleic acid construct encoding the fusion protein in the  
 host cell, purifying the fusion protein on a solid support capable of  
 specifically binding to the first purification tag, contacting the  
 purified fusion protein with an amino-terminal domain of the same Ubl  
 molecule comprising a second purification tag, thus generating a  
 reconstituted Ubl, purifying the reconstituted Ubl on a solid support  
 capable of specifically binding the second purification tag, and  
 contacting the purified reconstituted Ubl with a protease specific to the  
 Ubl molecule, thus cleaving the fusion protein such that the protein of  
 interest is now purified), increasing (M4) the affinity between a carboxy  
 -terminal domain of a Ubl molecule and a amino-terminal domain of a Ubl  
 molecule (involving operably linking a moiety to the carboxy-terminal and  
 to the amino-terminal domain, or inserting mutations into the carboxy-  
 terminal domain and the amino-terminal domain), a kit for purification of  
 a protein from a host cell, and generating an altered amino terminus in a  
 protein of interest in a host cell. The Ubl molecule is chosen from SUMO  
 (small ubiquitin related modifier), ubiquitin, RUB, HUB, APC8, APC12,  
 URM1, FAT10, Ubi-L and ISG15, preferably SUMO. The purification tag is  
 chosen from polyhistidine tag (6x His), polyarginine tag, glutathione-S-  
 transferase (GST), maltose binding protein (MBP), S-tag, influenza virus  
 HA tag, thioredoxin, staphylococcal protein A tag, the FLAG epitope,  
 AviTag epitope, and the c-myc epitope. The methods are useful for  
 enhancing the expression of a protein of interest in a host cell chosen  
 from yeast cell, Escherichia coli, bacterial cell, mammalian cell and  
 insect cell. The yeast SUMO N terminus is designated NTHS (N-terminal  
 half of SUMO) and the C-terminus CTHS. The present sequence represents a  
 ubiquitin like protein (or its N- or C-terminus) suitable for use in a  
 fusion protein of the invention.

Sequence 32 AA;

Alignment Scores:

Pred. No.: 7.81e-08 Length: 32  
 Score: 114.00 Matches: 21  
 Percent Similarity: 84.4% Conservative: 6  
 Best Local Similarity: 65.6% Mismatches: 5



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:39:29 ; Search time 1.62076 Seconds  
(without alignments)  
1211.048 Million cell updates/sec

Title: US-10-067-832D-13  
Perfect score: 188  
Sequence: 1 atgatcgaggttcttgcga.....agaagctgattgcagcctaa 102

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat 07042006 175613 9739/app query.fasta\_1  
-DB=PIR -OPMT=fastan -SUPFIK=n2p.rpr -MINMATCH=0.1 -LOOPECI=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -NATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DICALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USR=US10067832 @CGN 1.1.77 @runat 07042006 175613 9739 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	75.5	73	2 T47458	hypothetical prote
2	140	74.5	73	2 T40200	ubiquitin-like pro
3	135	71.8	73	2 T25763	hypothetical prote
4	114	60.6	73	2 S78735	protein YNR032c-a
5	56.5	30.1	282	2 S27679	NAD glycohydrolase
C 6	55.5	29.5	2695	2 S54974	type 1 inositol 1,
C 7	55.5	29.5	2713	2 A55713	inositol 1,4,5-tri
C 8	55.5	29.5	2734	2 B36579	inositol 1,4,5-tri
C 9	55.5	29.5	2749	1 ACMSIT	inositol 1,4,5-tri
C 10	55.5	29.5	2749	2 A36579	inositol 1,4,5-tri
C 11	54.5	29.0	2693	2 A40743	IP3 receptor, XIP3
C 12	52.5	27.9	831	2 A70363	mannose-1-phosphat
C 13	52.5	27.9	1170	2 S52525	probable membrane
C 14	52.5	27.9	2701	2 S17796	inositol-trisphosph

C 15	52	27.7	180	2 AC0197	conserved hypothet
C 16	52	27.7	371	2 A71359	conserved hypothet
C 17	52	27.7	1587	2 AB2012	hypothetical prote
C 18	52	27.7	1601	2 AB1730	hypothetical prote
C 19	52	27.7	4543	1 A53102	alpha-2-macroglobu
C 20	51.5	27.4	201	2 T31652	hypothetical prote
C 21	51	27.1	410	2 F96683	hypothetical prote
C 22	51	27.1	527	2 S42512	recombination-acti
C 23	51	27.1	662	2 T01857	hypothetical prote
C 24	51	27.1	3131	2 T39553	vacuolar protein a
C 25	51	27.1	4544	1 S02392	alpha-2-macroglobu
C 26	51	27.1	4545	1 S25111	alpha-2-macroglobu
C 27	50.5	26.9	1290	2 A55094	chromosomal protei
C 28	50	26.6	69	2 A25161	tray protein - Sal
C 29	50	26.6	88	2 JC5203	outer membrane pro
C 30	50	26.6	90	2 B86560	9 kDa-Cysteine-ric
C 31	50	26.6	90	2 A72064	cysteine rich oute
C 32	50	26.6	309	2 S66265	L-ascorbate peroxi
C 33	50	26.6	355	2 B97060	uncharacterized pr
C 34	50	26.6	459	2 T37677	hypothetical integ
C 35	50	26.6	525	2 B81859	phospholipase D-fa
C 36	50	26.6	968	2 C64691	type III restricti
C 37	50	26.6	972	2 C71826	hypothetical prote
C 38	49.5	26.3	498	2 AD0854	probable permease
C 39	49.5	26.3	829	2 S12858	vira protein - Agr
C 40	49.5	26.3	1209	2 T31657	reverse transcript
C 41	49.5	26.3	1965	2 S75200	fat protein - Syne
C 42	49	26.1	392	2 AE3382	nifs protein (impo
C 43	49	26.1	434	2 T49044	hypothetical prote
C 44	49	26.1	508	2 B81083	cardiolipin synthe
C 45	49	26.1	551	2 C84549	probable ubiquitin

ALIGNMENTS

RESULT 1

T47458 hypothetical protein T14D3.120 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C/Accession: T47458

R/Jordan, N.J. Bangert, S.; Wiedelmann, R.; Voss, H.; Unseeld, M.; Mewes, H.W.; Lemcke, K.

Submitted to the Protein Sequence Database, February 2000

A/Reference number: Z24467

A/Accession: T47458

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-73 <JOR>

A/Cross-references: UNIPROT:Q9M1U1; UNIPARC:UPI000000A81D7; EMBL:AL138649

A/Experimental source: cultivar Columbia; BAC clone T14D3

C/Genetics:

A/Map position: 3

A/Note: T14D3.120

Alignment Scores:	1.48e-12	Length:	73
Pred. No.:	142.00	Matches:	28
Score:	90.9%	Conservative:	2
Percent Similarity:	84.8%	Mismatches:	3
Best Local Similarity:	75.5%	Indels:	0
Query Match:	2	Gaps:	0
DB:			

US-10-067-832D-13 (1-102) x T47458 (1-73)

QY 1 ATGATCGAGGTCTGTTTGGCAACGACCTCTGGGAAAAAGGTCCTCGTTAAATGCAACACG 60

|||||

Db 1 MetilcgluValValLeuAnaAspArgLeuGlyLysValArgValLysCysAsnGlu 20

QY 61 GATGATACCATCGGGACCTTAAAGAGCTGATTGCAGCC 99

|||||

Db 21 GluAspThrIleGlyAspLeuLysLeuValAlaAala 33

RESULT 2

T40200  
ubiquitin-like protein - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T40200  
R;Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: Z21912  
A;Accession: T40200  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-73 <OLI>  
A;Cross-references: UNIPROT:O94650; UNIPARC:UPI000006A678; EMBL:AL049190; PIDN:CAB39137.  
A;Experimental source: strain 972h-; cosmid C31E1  
C;Genetics:  
A;Gene: SPDB:SPBC31E1.03  
A;Map position: 2  
A;Introns: 6/2; 40/2

Alignment Scores:  
Pred. No.: 2,94e-12 Length: 73  
Score: 140.00 Matches: 26  
Percent Similarity: 87.9% Conservative: 3  
Best Local Similarity: 78.8% Mismatches: 4  
Query Match: 74.5% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x T40200 (1-73)

QY 1 ATGATCGAGGTGTTTGCACGACCGCTCTGGGAAAAGGTCNCGTAAATGCAACACG 60  
|||||  
Db 1 MetileGluValLeuGlyAspArgLeuGlyValArgValGlyValGlyCysMetPro 20  
|||||

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrValGlyAspPheLysLeuValAlaAla 33  
|||||

RESULT 3  
T25763  
hypothetical protein F46F11.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T25763  
R;Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, February 1997  
A;Description: The sequence of C. elegans cosmid F46F11.  
A;Reference number: Z20083  
A;Accession: T25763  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-73 <PAU>  
A;Cross-references: UNIPROT:P91302; UNIPARC:UPI00000076B0C; EMBL:U88173; PIDN:AAB42266.1  
A;Experimental source: strain Bristol N2; clone F46F11  
C;Genetics:  
A;Gene: CESP:F46F11.4  
A;Map position: 1  
A;Introns: 38/2

Alignment Scores:  
Pred. No.: 1.62e-11 Length: 73  
Score: 135.00 Matches: 26  
Percent Similarity: 84.8% Conservative: 2  
Best Local Similarity: 78.8% Mismatches: 5  
Query Match: 71.8% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x T25763 (1-73)

QY 1 ATGATCGAGGTGTTTGCACGACCGCTCTGGGAAAAGGTCNCGTAAATGCAACACG 60  
|||||  
Db 1 MetileGluIleThrValAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20  
|||||

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99

Db 21 SerAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
|||||

RESULT 4  
S78735  
protein YNR032c-a - Yeast (Saccharomyces cerevisiae)  
C;Species: Saccharomyces cerevisiae  
C;Date: 15-Jan-1999 #sequence\_revision 15-Jan-1999 #text\_change 15-Jan-1999  
C;Accession: S78735  
R;Pohl, T.M.  
submitted to the Protein Sequence Database, April 1996  
A;Reference number: S63346  
A;Accession: S78735  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-73 <POH>  
A;Cross-references: UNIPARC:UPI0000168403; EMBL:Z71647; MIPS:YNR032c-a  
A;Experimental source: strain S288C  
C;Genetics:  
A;Map position: 14R

Alignment Scores:  
Pred. No.: 2.13e-08 Length: 73  
Score: 114.00 Matches: 21  
Percent Similarity: 84.4% Conservative: 6  
Best Local Similarity: 65.6% Mismatches: 5  
Query Match: 60.6% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x S78735 (1-73)

QY 1 ATGATCGAGGTGTTTGCACGACCGCTCTGGGAAAAGGTCNCGTAAATGCAACACG 60  
|||||  
Db 1 MetileGluValValValAsnAspArgLeuGlyLysValArgValLysCysLeuAla 20  
|||||

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCA 96  
|||||  
Db 21 GluAspSerValGlyAspPheLysLysValLeuSer 32  
|||||

RESULT 5  
S27769  
NAD glycohydrolase - California sea hare  
C;Species: Aplysia californica (California sea hare)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: S27769  
R;Glück, D.L.; Helmich, M.R.; Beushausen, S.; Tempst, P.J.; Bayley, H.; Strumwasser, F.  
submitted to the EMBL Data Library, March 1992  
A;Description: Primary structure and localization of a NAD glycohydrolase from Aplysia c.  
A;Reference number: S27769  
A;Accession: S27769  
A;Molecule type: mRNA  
A;Residues: 1-282 <GLI>  
A;Cross-references: UNIPROT:P29241; UNIPARC:UPI000012FC50; EMBL:M85206; NID:G155768; PID:  
C;Superfamily: ADP-ribosyl cyclase CD38  
C;Keywords: NAD

Alignment Scores:  
Pred. No.: 6.8 Length: 282  
Score: 56.50 Matches: 12  
Percent Similarity: 62.5% Conservative: 8  
Best Local Similarity: 37.5% Mismatches: 11  
Query Match: 30.1% Indels: 1  
DB: 2 Gaps: 1

US-10-067-832D-13 (1-102) x S27769 (1-282)

QY 4 ATCGAGGTGTTTGCACGACCGCTCTGGGAAAAGGTCNCGTAAATGCAACACGAT 63  
|||||  
Db 214 ValLysValIleValLeuHisArgLeuGlyGlyLysIleIleGlyLysCysGlyAlaGly 233  
|||||

QY 64 GATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
Db 234 SerLeuLeu---AspLeuGluLysLeuValLysAla 244  
|||||



C;Date: 03-Mar-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004



N;Alternate names: hypothetical protein LPa11w; hypothetical protein YP9132.07  
C;Species: Saccharomyces cerevisiae  
C;Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C;Accession: S52525; S59687  
R;Badcock, K.; Churcher, C.  
submitted to the EMBL Data Library, February 1995  
A;Reference number: S52519  
A;Accession: S52525  
A;Molecule type: DNA  
A;Residues: 1-1170 <BAD>  
A;Cross-references: UNIPROT:Q12200; UNIPARC:UPI000006A19C; EMBL:Z48483; NID:G683777; PID:1170  
A;Experimental source: strain AB972  
R;Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa  
submitted to the EMBL Data Library, August 1995  
A;Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.  
A;Reference number: S59677  
A;Accession: S59687  
A;Molecule type: DNA  
A;Residues: 1-1170 <HAL>  
A;Cross-references: UNIPARC:UPI000006A19C; EMBL:U33335; NID:G965076; PID:G965087; MIPS:Y  
C;Genetics:  
A;Gene: SGD:NCR1  
A;Cross-references: SGD:S0005927; MIPS:YPL006w  
A;Map position: 1cL  
C;Keywords: transmembrane protein  
F;1-17/Domain: transmembrane #status predicted <TM1>  
F;258-274/Domain: transmembrane #status predicted <TM2>  
F;341-357/Domain: transmembrane #status predicted <TM3>  
F;585-601/Domain: transmembrane #status predicted <TM4>  
F;614-630/Domain: transmembrane #status predicted <TM5>  
F;667-683/Domain: transmembrane #status predicted <TM6>  
F;698-714/Domain: transmembrane #status predicted <TM7>  
F;751-767/Domain: transmembrane #status predicted <TM8>  
F;1004-1020/Domain: transmembrane #status predicted <TM9>  
F;1027-1043/Domain: transmembrane #status predicted <TM10>  
F;1051-1067/Domain: transmembrane #status predicted <TM11>  
F;1103-1119/Domain: transmembrane #status predicted <TM12>  
F;1137-1153/Domain: transmembrane #status predicted <TM13>  
  
Alignment Scores:  
Pred. No.: 24.6 Length: 1170  
Score: 52.50 Matches: 11  
Percent Similarity: 70.4% Conservatives: 8  
Best Local Similarity: 40.7% Mismatches: 7  
Query Match: 27.9% Indels: 1  
DB: 2 Gaps: 1  
  
US-10-067-832D-13 (1-102) x S52525 (1-1170)  
  
QY 84 CTTAAGTCCCGATGATATCCGTTGTCATTTAACGNGGACCTTTTCCCGACAGC 25  
|||||  
Db 532 LeuLysValProGluGly---LeuArgIleSerPheAsnThrGluIleSerLeuGluLys 550  
:::  
  
QY 24 GTCGTTCAACACCTCGAT 4  
:::  
Db 551 GluLeuAsnAsnAsnAsnAsnAsp 557  
:::  
  
RESULT 14  
S17796  
inositol-trisphosphate receptor type 2 - rat  
N;Alternate names: inositol-1,4,5-trisphosphate receptor type 2  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S17796  
R;Suedhof, T.C.; Newton, C.L.; Archer III, B.T.; Ushkaryov, Y.A.; Mignery, G.A.  
EMBO J. 10, 3199-3206, 1991  
A;Title: Structure of a novel Insp(3) receptor.  
A;Reference number: S17796; NUID:92007769; PMID:1655411  
A;Accession: S17796  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-2701 <SUE>  
A;Cross-references: UNIPROT:P29995; UNIPARC:UPI000012D7CA; EMBL:X61677; NID:G56507; PIDN

C;Superfamily: inositol-trisphosphate receptor  
C;Keywords: transmembrane protein

Alignment Scores:  
Pred. No.: 23.4 Length: 2701  
Score: 52.50 Matches: 13  
Percent Similarity: 58.8% Conservatives: 7  
Best Local Similarity: 38.2% Mismatches: 11  
Query Match: 27.9% Indels: 3  
DB: 2 Gaps: 2

US-10-067-832D-13 (1-102) x S17796 (1-2701)

QY 96 TGCATCAGCTTCTTAAGTCCCGATGTCATCATCGTTGTCATTTAACGNGGACCTT 37  
|||||  
Db 1397 CysAsnSerLeuLeuProLeu---AspAspIleValArgValThrHisAspAspCys 1415  
:::

QY 36 TTTCCCGACAGCGTCGTGCA-----AACACCTCGATCAT 1  
|||||

Db 1416 IleProGluValIleAlaTyValAsnPheValAsnHis 1429  
:::

#### RESULT 15

AC0197

Conserved hypothetical protein YP01616 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AC0197

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; NUID:21470413; PMID:11586360  
A;Accession: AC0197  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-180 <KUR>  
A;Cross-references: UNIPROT:Q8ZFS2; UNIPARC:UPI000013AF21; GB:AL590842; PIDN:CAC90438.1;  
C;Genetics:  
A;Gene: YP01616

C;Superfamily: Escherichia coli ycfp protein

Alignment Scores:

Pred. No.: 32.5 Length: 180  
Score: 52.00 Matches: 10  
Percent Similarity: 72.7% Conservatives: 6  
Best Local Similarity: 45.5% Mismatches: 6  
Query Match: 27.7% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x AC0197 (1-180)

QY 69 GGTATCATCGCTGTCGTTTACGNGGACCTTTTCCCGACAGCGTTCGTTGCAACAC 10  
|||||  
Db 82 GlyIleArgGlnValAlaPheAsnProAsnLeuTyProGlnGluAsnMetSerGlyLys 101  
:::

QY 9 CTCGAT 4  
:::

Db 102 IleAsp 103

Search completed: April 7, 2006, 23:47:49  
Job time : 15.1038 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:32:23 ; Search time 9.42203 Seconds  
(without alignments)  
1527.566 Million cell updates/sec

Title: US-10-067-832D-13  
Perfect score: 188  
Sequence: 1 atgatcgaggtgtttgtcaa.....agaagtgtatgcagcctaa 102

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-O=/abs/ABSSWEB.spool/US10067832/runat\_07042006\_175611\_9700/app\_query.fasta\_1  
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USBR=US10067832 @CGN 1.1.580 @runat\_07042006\_175611\_9700 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	86.7	47	2 Q5BJS1_RAT	Q5bjs1 rattus norv
2	163	86.7	73	1 UBL5_HUMAN	Q9bz11 homo sapien
3	163	86.7	73	1 UBL5_MESAU	Q9egx7 mesocricetu
4	163	86.7	73	1 UBL5_MOUSE	Q9epv8 mus musculu
5	163	86.7	73	1 UBL5_PSAOB	Q791b0 psammomya o
6	163	86.7	73	2 Q5RC10_PONPY	Q5rc10 pongo pygma
7	163	86.7	73	2 Q4RS53_MACFA	Q4rs53 macaca fasc
8	153	81.4	73	1 UBL5_BRARE	Q7sx12 brachydanio
9	152	80.9	73	1 UBL5_DROME	Q9v998 drosophila
10	151	80.3	73	2 Q5H2C1_XENLA	Q5h2c1 xenopus lae
11	150	79.8	73	2 Q7QHM2_ANOGA	Q7qhm2 anopheles g
12	147	78.2	73	2 Q4SMI6_TETNG	Q4smi6 tetraodon n
13	146	77.7	73	1 UBL5_ARATH	Q9fg29 arabidopsis
14	146	77.7	73	2 Q6K220_ORYSA	Q6k220 oryza sativ
15	146	77.7	73	2 Q570V8_ARATH	Q570v8 arabidopsis
16	143	76.1	73	2 Q81545_PLAF7	Q81545 plasmodium

17	143	76.1	73	2 Q7RBW4_PLAYO	Q7rbw4 plasmodium
18	143	76.1	73	2 Q4YUJ7_PLABE	Q4yuj7 plasmodium
19	142	75.5	73	2 Q9MIU1_ARATH	Q9miu1 arabidopsis
20	140	74.5	73	1 HUB1_SCHPO	O94650 schizosacch
21	136	72.3	87	2 Q540Q3_DICDI	O94603 dictyosteli
22	135	71.8	73	1 UBL5_CAEEL	P91302 caenorhabdi
23	135	71.8	73	2 Q617M9_CAEBR	Q617m9 caenorhabdi
24	133	70.7	73	2 Q4P9W2_USTWA	Q4p9w2 ustilago ma
25	133	70.7	73	2 Q4MZ69_THEPA	Q4mz69 theileria p
26	133	70.7	76	2 Q4UBE9_THEAN	Q4ube9 theileria a
27	131	69.7	73	2 Q5BS57_SCHJA	Q5bs57 schistosoma
28	130	69.1	70	2 Q6CT04_YARLI	Q6ct04 yarrowia li
29	128	68.1	73	2 Q4VAF8_MOUSE	Q4vaf8 mus musculu
30	121	64.4	276	2 Q5SM54_CRYNE	Q5sm54 cryptococcu
31	121	64.4	276	2 Q5K815_CRYNE	Q5k815 cryptococcu
32	120	63.8	73	1 HUB1_DEBHA	O8up7 debaryomyce
33	117	62.2	73	1 HUB1_CANGA	O8fix7 candida gla
34	116	61.7	74	1 HUB1_KLULA	Q8cul2 kluyveromyc
35	116	61.7	79	2 Q6CUI1_KLULA	Q6cul1 kluyveromyc
36	115	61.2	73	1 HUB1_ASHGO	Q756x3 ashbya gos
37	110	58.5	73	1 HUB1_YEAST	Q6g546 saccharomyc
38	96	51.1	79	2 Q50P03_ENTHI	Q50p03 entamoeba h
39	86	45.7	616	2 Q5BCG2_EMENI	Q5bcg2 aspergillus
40	85	45.2	237	2 Q41016_GIBZE	Q41016 gibberella
41	69	36.7	239	2 Q52620_MAGGR	Q52620 magnaporthe
C 42	66	35.1	333	2 Q8A116_BACTN	Q8a116 bacteroides
C 43	64.5	34.3	1417	2 Q4S7P9_TETNG	Q4s7p9 tetraodon n
C 44	62	33.0	189	2 Q4WK66_ASPFU	Q4wk66 aspergillus
45	62	33.0	261	2 Q8X065_NEUCR	Q8x065 neurospora

ALIGNMENTS

RESULT 1  
Q5BJS1\_RAT  
ID Q5BJS1\_RAT PRELIMINARY; PRT; 47 AA.  
AC Q5BJS1;  
DT 10-MAY-2005 (TREMBlrel. 30, Created)  
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
DE 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE LOC500954 protein.  
GN Names:LOC500954;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.



```

OX NCBI_TaxID=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA "Beacon in the golden hamster."
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL -1- SUBUNIT: Interacts with CLK4 (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AY329082; AAQ99044.1; -; mRNA.
DR SMR; Q6EGX7; 1-73.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PROSITE; PS00299; UBIQUITIN_1; FALSE NEG.
DR PROSITE; PS50053; UBIQUITIN_2; FALSE NEG.
KW UBL conjugation pathway.
FT DOMAIN 1 73 Ubiquitin-like.
SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:
Pred. No.: 4.3e-16 Length: 73
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 86.7% Indels: 0
DB: 1 Gaps: 0

US-10-067-832D-13 (1-102) x UBL5_MESAU (1-73)
QY 1 ATGATCGAGGTGTTTCACAGCCGCTCGGGAAAGGTCNCGTTAATGCAACG 60
DB 1 MetteliuValValCyAsnAspArgLeuGlyLysValArgValLysCyAsnThr 20

QY 61 GATGATACCATCGGGACCTTAAGAACGTGATTCAGCC 99
DB 21 AspAspThrileGlyAspLeuLysLeuIleAla 33

RESULT 4
UBL5_MOUSE
ID UBL5_MOUSE STANDARD; PRT; 73 AA.
AC Q9EPV8;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ubiquitin-like protein 5.
GN Name=UBL5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Kidney, Pancreas, Spinal ganglion, and Tongue;
RA MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldrinelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Glessi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

```

```

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yangisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Colling F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S., Krzywinski M.I., Skalek U., Smalls D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBUNIT: Interacts with CLK4 (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AK002730; BAB22312.1; -; mRNA.
DR EMBL; AK003992; BAB23111.1; -; mRNA.
DR EMBL; AK007726; BAB25215.1; -; mRNA.
DR EMBL; AK009854; BAB26545.1; -; mRNA.
DR EMBL; AK012803; BAB28481.1; -; mRNA.
DR EMBL; AK051149; BAC34537.1; -; mRNA.
DR EMBL; BC028498; AAB28498.1; -; mRNA.
DR PDB; 1UH6; NMR; A=1-73.
DR MGI; MGI:1913427; Ubl5.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PROSITE; PS00299; UBIQUITIN_1; FALSE NEG.
DR PROSITE; PS50053; UBIQUITIN_2; FALSE NEG.
KW 3D-structure; Ubl conjugation pathway.
FT DOMAIN 1 73 Ubiquitin-like.
SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:
Pred. No.: 4.3e-16 Length: 73
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0

```

Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 1 Gaps: 0

US-10-067-832D-13 (1-102) x UBL5\_MOUSE (1-73)

QY 1 ATGATCGAGGTGTTTTCACACGACCGTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
|||||  
DB 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
|||||  
QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
|||||

## RESULT 5

UBL5\_PSAOB  
ID UBL5\_PSAOB STANDARD; PRT; 73 AA.  
AC Q791B0;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Ubiquitin-like protein 5 (Beacon protein).  
GN Names=UBL5;  
OS Psammomys obesus (Fat sand rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Gerbillinae; Psammomys.  
OX NCBI\_TaxID=48139;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Hypothalamus;  
RX MEDLINE=20527879; PubMed=11078442;  
RA Collier G.R., McMillan J.S., Windmill K., Walder K., Tenne-Brown J.,  
RA de Silva A., Trevaskis J., Jones S., Morton G.J., Lee S., Augert G.,  
RA Civitarese A., Zimet P.Z.;  
RT "Beacon: a novel gene involved in the regulation of energy balance.";  
RL Diabetes 49:1766-1771(2000).  
CC -|- SUBUNIT: Interacts with CLK4 (By similarity).  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -|- SIMILARITY: Contains 1 ubiquitin-like domain.

-----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC EMBL; AF318186; AAC34704.1; -; mRNA.  
DR SMR; Q791B0; 1-73.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin.1.  
DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
DR PROSITE; PS50053; UBIQUITIN\_2; FALSE\_NEG.  
KW Ub1 conjugation pathway.  
FT DOMAIN 1 73 Ubiquitin-like.  
SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

Alignment Scores:  
Pred. No.: 4.3e-16 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 1  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 1 Gaps: 0

US-10-067-832D-13 (1-102) x UBL5\_PSAOB (1-73)

QY 1 ATGATCGAGGTGTTTTCACACGACCGTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
|||||  
DB 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
|||||  
QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||

Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

## RESULT 6

QSRC10\_PONPY  
ID QSRC10\_PONPY PRELIMINARY; PRT; 73 AA.  
AC QSRC10;  
DT 01-FEB-2005 (TREMBLrel. 29, Created)  
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)  
DE Hypothetical protein DKFPz469G145.  
GN Name=DKFPz469G145;  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RG The German cDNA Consortium;  
RA Ansorge W., Kleiger S., Regiert T., Rittmuller C., Schwager B.,  
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CR858290; CAH90527.1; -; mRNA.  
DR SMR; QSRC10; 1-73.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin.1.  
KW Hypothetical protein.

SQ SEQUENCE 73 AA; 8547 MW; D68BC941536AA39F CRC64;

## Alignment Scores:

Pred. No.: 4.3e-16 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x QSRC10\_PONPY (1-73)

QY 1 ATGATCGAGGTGTTTTCACACGACCGTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
|||||  
DB 1 MetileGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
|||||  
QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
|||||

## RESULT 7

Q4R5J3\_MACFA  
ID Q4R5J3\_MACFA PRELIMINARY; PRT; 73 AA.  
AC Q4R5J3;  
DT 13-SEP-2005 (TREMBLrel. 31, Created)  
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
DE Brain cDNA, clone: QFLA-12296, similar to human ubiquitin-like 5  
DE (UBL5), (Testis cDNA, clone: QtsA-15893, similar to human ubiquitin-  
DE like 5 (UBL5)).  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
OC Cercopitheidae; Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RA International consortium for macaque cDNA sequencing, analysis;  
RT "DNA sequences of macaque genes expressed in brain or testis and its  
RT evolutionary implications."  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RA Oeada N., Hirata M., Tanuma R., Kueada J., Hida M., Suzuki Y.,

RP  
RX MEDLINE=22426069; PubMed=12537572;  
RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,





RG The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008816; EAA05206.2; -; Genomic\_DNA.  
 DR SMR; O7QHM2; 1-73.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PROSITE; PS0053; UBIQUITIN 2; 1.  
 SQ SEQUENCE 73 AA; 8599 MW; 43D212BE601B3COD CRC64;

Alignment Scores:  
 Pred. No.: 4,69e-14 Length: 73  
 Score: 150.00 Matches: 28  
 Percent Similarity: 90.9% Conservative: 2  
 Best Local Similarity: 84.8% Mismatches: 3  
 Query Match: 79.8% Indels: 0  
 DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x Q7QHM2\_ANOGA (1-73)  
 QY 1 ATGATCGAGGTGTTTGCACACGCGTCTGGGAAAAGTCCNCGTTAAATGCAACACG 60  
 DB 1 MetLeuGluIleThrCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnPro 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTCAGCC 99  
 DB 21 AspAspThrIleGlyAspLeuLysLysLeuIleAla 33

RESULT 12  
 Q4SM16\_TETNG  
 ID Q4SM16\_TETNG PRELIMINARY; PRT; 73 AA.  
 AC Q4SM16;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Chromosome 18 SCAF14547, whole genome shotgun sequence.  
 DE (Fragment).  
 OS ORFNames=GSTENG00015743001;  
 GN Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OC NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Blomont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,  
 RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype."  
 RL Nature 431:946-957(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAE01014547; CAF98146.1; -; Genomic\_DNA.  
 FT NON TER 73  
 SQ SEQUENCE 73 AA; 8592 MW; 6A14CA0323F8E8EB CRC64;

Alignment Scores:  
 Pred. No.: 1,39e-13 Length: 73  
 Score: 147.00 Matches: 29  
 Percent Similarity: 90.9% Conservative: 1  
 Best Local Similarity: 87.9% Mismatches: 3  
 Query Match: 78.2% Indels: 0  
 DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x Q4SM16\_TETNG (1-73)  
 QY 1 ATGATCGAGGTGTTTGCACACGCGTCTGGGAAAAGTCCNCGTTAAATGCAACACG 60  
 DB 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnPro 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTCAGCC 99  
 DB 21 GlnAspSerIleGlyAspLeuLysLysLeuIleAla 33

RESULT 13  
 UBL5\_ARATH  
 ID UBL5\_ARATH STANDARD; PRT; 73 AA.  
 AC Q9PGZ9;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Ubiquitin-like protein 5.  
 GN Name=UBL5; OrderedLocName=At5g42300; ORFNames=K5J14.10;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjani M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koseuma E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamiya A., Meyers R., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tanase R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome."  
 RL Science 302:842-846(2003).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL; AB023032; BAB10203.1; -; Genomic\_DNA.  
DR EMBL; AY072534; AAL66949.1; -; mRNA.  
DR EMBL; AF370527; AAK48954.1; -; mRNA.  
DR SMR; Q9FGZ9; 1-72.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin; 1.  
DR PROSITE; PS00299; UBIQUITIN\_1; FALSE\_NEG.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
KW Ubl conjugation pathway.  
FT DOMAIN 1 73 Ubiquitin-like.  
SQ SEQUENCE 73 AA; 8571 MW; 4F20C6ED077133F6 CRC64;

Alignment Scores:  
Pred. No.: 1.99e-13 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 1 Gaps: 0

US-10-067-832D-13 (1-102) x UBL5\_ARATH (1-73)

QY 1 ATGATCGAGTGTGTTGCAACGACCGTCTGGGAAAAGTCNCGTTAAATGCAACACG 60  
|||||  
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgVallysCysAsnAsp 20  
QY 61 GATGATACCATCGGGACCTTAAGAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuValAlaAa 33

RESULT 14  
Q6K220\_ORYSA  
ID Q6K220\_ORYSA PRELIMINARY; PRT; 73 AA.  
AC Q6K220;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Ubiquitin-like protein.  
GN Name=B146H02.16;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Katayose Y.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP061168; BAD23730.1; -; Genomic\_DNA.  
DR SMR; Q6K220; 1-72.  
DR Gramene; Q6K220; -.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin; 1.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 73 AA; 8529 MW; FDA1C6ED0770F525 CRC64;

Alignment Scores:  
Pred. No.: 1.99e-13 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x Q6K220\_ORYSA (1-73)

QY 1 ATGATCGAGTGTGTTGCAACGACCGTCTGGGAAAAGTCNCGTTAAATGCAACACG 60

Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgVallysCysAsnGlu 20  
QY 61 GATGATACCATCGGGACCTTAAGAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuValAlaAa 33

RESULT 15

Q570V8\_ARATH  
ID Q570V8\_ARATH PRELIMINARY; PRT; 73 AA.  
AC Q570V8;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Ubiquitin-like protein.  
GN Name=At5g42300;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,  
RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,  
RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,  
RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,  
RA Hayashizaki Y., Shinozaki K.;  
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK220599; BAD94931.1; -; mRNA.  
SQ SEQUENCE 73 AA; 8571 MW; 4F20C6ED077133F6 CRC64;

Alignment Scores:

Pred. No.: 1.99e-13 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x Q570V8\_ARATH (1-73)

QY 1 ATGATCGAGTGTGTTGCAACGACCGTCTGGGAAAAGTCNCGTTAAATGCAACACG 60  
|||||  
Db 1 MetileGluValValLeuAenAspArgLeuGlyLysValArgVallysCysAsnAsp 20  
QY 61 GATGATACCATCGGGACCTTAAGAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuValAlaAa 33

Search completed: April 7, 2006, 23:46:27

Job time : 52.1102 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:46:49 ; Search time 2.16102 Seconds  
(without alignments)  
780.458 Million cell updates/sec

Title: US-10-067-832D-13  
Perfect score: 188  
Sequence: 1 atgatcgaggttgttgcga.....agaagctgattgcgcctaa 102

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832/runat 07042006 175615 9797/app query.fasta\_1  
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosu62 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs802p -USER=US10067832 @CGN 1.1.101 @runat 07042006 175615 9797  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	86.7	33	2	US-09-331-930A-14
2	163	86.7	73	2	US-09-331-930A-2
3	163	86.7	73	2	US-09-331-930A-19
4	163	86.7	73	2	US-09-331-930A-20
5	163	86.7	73	2	US-09-331-930A-21
6	163	86.7	73	2	US-09-331-930A-24
7	163	86.7	73	2	US-09-331-930A-25
8	135	71.8	73	2	US-09-331-930A-22
9	106	56.4	66	2	US-09-331-930A-23
10	106	56.4	75	2	US-09-248-796A-16302
11	59	31.4	92	2	US-09-252-991A-29573
12	58	30.9	644	2	US-09-248-796A-17325

c 13	55.5	29.5	2713	4	PCT-US96-01735-1	Sequence 1, Appli
c 14	55.5	29.5	2749	2	US-09-385-222A-4	Sequence 4, Appli
c 15	54.5	29.0	456	2	US-09-328-352-6312	Sequence 6312, Ap
c 16	52.5	27.9	213	2	US-10-104-047-2722	Sequence 2722, Ap
c 17	52.5	27.9	1170	2	US-09-462-136-6	Sequence 6, Appli
c 18	52.5	27.9	2262	2	US-09-949-016-8849	Sequence 8849, Ap
c 19	52	27.7	222	2	US-09-252-991A-21885	Sequence 21885, A
c 20	52	27.7	231	2	US-09-107-532A-5301	Sequence 5301, Ap
c 21	52	27.7	341	2	US-09-252-991A-32424	Sequence 32424, A
c 22	52	27.7	1094	2	US-09-268-347-32	Sequence 32, Appl
c 23	52	27.7	1761	2	US-09-489-039A-11234	Sequence 11234, A
c 24	51.5	27.4	487	2	US-10-012-231A-63	Sequence 63, Appl
c 25	51.5	27.4	487	2	US-10-015-389A-63	Sequence 63, Appl
c 26	51.5	27.4	487	2	US-10-006-768A-63	Sequence 63, Appl
c 27	51.5	27.4	487	2	US-10-015-671A-63	Sequence 63, Appl
c 28	51.5	27.4	487	2	US-10-015-393A-63	Sequence 63, Appl
c 29	51.5	27.4	487	2	US-10-011-893A-63	Sequence 63, Appl
c 30	51.5	27.4	487	2	US-10-006-041A-63	Sequence 63, Appl
c 31	51.5	27.4	487	2	US-10-012-064A-63	Sequence 63, Appl
c 32	51	27.1	4544	1	US-08-469-486-52	Sequence 52, Appl
c 33	51	27.1	4544	1	US-08-469-658-52	Sequence 52, Appl
c 34	50	26.6	90	2	US-09-198-452A-597	Sequence 597, App
c 35	50	26.6	92	2	US-09-438-185A-560	Sequence 560, App
c 36	50	26.6	221	2	US-09-270-767-48970	Sequence 48970, A
c 37	50	26.6	482	2	US-09-270-767-43232	Sequence 43232, A
c 38	50	26.6	512	2	US-09-270-767-42979	Sequence 42979, A
c 39	50	26.6	921	2	US-09-902-540-15467	Sequence 15467, A
c 40	49	26.1	113	2	US-09-173-300-54	Sequence 54, Appl
c 41	49	26.1	113	2	US-10-027-450-54	Sequence 54, Appl
c 42	49	26.1	302	2	US-09-270-767-44347	Sequence 44347, A
c 43	49	26.1	347	2	US-09-270-767-57323	Sequence 57323, A
c 44	49	26.1	348	2	US-09-248-796A-18141	Sequence 18141, A
c 45	49	26.1	349	2	US-09-270-767-42065	Sequence 42065, A

ALIGNMENTS

RESULT 1  
US-09-331-930A-14  
; Sequence 14, Application US/09331930A  
; Patent No. 6436670  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMER, PAUL Z.  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCES: 22975-20007.00  
; CURRENT APPLICATION NUMBER: US/09/331,930A  
; CURRENT FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP0117/97  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP0323/97  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (15)  
; OTHER INFORMATION: variable amino acid  
US-09-331-930A-14  
  
Alignment Scores:  
Pred. No.: 4.25e-17 Length: 33  
Score: 163.00 Matches: 33  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 86.7% Indels: 0  
DB: 2 Gaps: 0

```
US-10-067-832D-13 (1-102) x US-09-331-930A-14 (1-33)
QY 1 ATGATCGAGTGTCTTTGCAACGACCGTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
  |||
Db 1 MetIleGluValValCysAenAspArgLeuGlyLysLysVal***VallysCysAsnThr 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99
  |||
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

RESULT 2
US-09-331-930A-2
; Sequence 2, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid
; OTHER INFORMATION: sequence for beacon from unknown organism
US-09-331-930A-2
Alignment Scores:
Pred. No.: 4.8e-17 Length: 73
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 86.7% Indels: 0
Gaps: 2
US-10-067-832D-13 (1-102) x US-09-331-930A-2 (1-73)
QY 1 ATGATCGAGTGTCTTTGCAACGACCGTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
  |||
Db 1 MetIleGluValValCysAenAspArgLeuGlyLysLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99
  |||
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

RESULT 3
US-09-331-930A-19
; Sequence 19, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
```

```
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid
; OTHER INFORMATION: sequence for beacon from unknown organism
US-09-331-930A-19
Alignment Scores:
Pred. No.: 4.8e-17 Length: 73
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 86.7% Indels: 0
Gaps: 2
US-10-067-832D-13 (1-102) x US-09-331-930A-19 (1-73)
QY 1 ATGATCGAGTGTCTTTGCAACGACCGTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
  |||
Db 1 MetIleGluValValCysAenAspArgLeuGlyLysLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99
  |||
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

RESULT 4
US-09-331-930A-20
; Sequence 20, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-331-930A-20
Alignment Scores:
Pred. No.: 4.8e-17 Length: 73
Score: 163.00 Matches: 32
Percent Similarity: 97.0% Conservative: 0
Best Local Similarity: 97.0% Mismatches: 1
Query Match: 86.7% Indels: 0
Gaps: 2
US-10-067-832D-13 (1-102) x US-09-331-930A-20 (1-73)
QY 1 ATGATCGAGTGTCTTTGCAACGACCGTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
  |||
Db 1 MetIleGluValValCysAenAspArgLeuGlyLysLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99
  |||
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33
```

```

Query Match:      86.7%      Indels:      0
DB:              2          Gaps:          0

US-10-067-832D-13 (1-102) x US-09-331-930A-24 (1-73)

Qy  1  ATGATCAGGTTGTTTGGCAACACCGCTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
Db  1  MetIleGluValValCysAsnAepArgLeuGlyLysLysValArgValLysCysAsnThr 20

Qy  61  GATGATACCATCGGGACCTTAAAGAGCTGATTGCAGCC 99
Db  21  AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

RESULT 7
US-09-331-930A-25
; Sequence 25, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007, 00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-331-930A-25

Alignment Scores:
Pred. No.:      4,8e-17      Length:      73
Score:          163.00      Matches:    32
Percent Similarity: 97.0%      Conservative: 0
Best Local Similarity: 97.0%      Mismatches:  1
Query Match:    86.7%      Indels:      0
DB:            2          Gaps:          0

US-10-067-832D-13 (1-102) x US-09-331-930A-25 (1-73)

Qy  1  ATGATCAGGTTGTTTGGCAACACCGCTCTGGGGAAAAAGGTCNCGTTAAATGCAACACG 60
Db  1  MetIleGluValValCysAsnAepArgLeuGlyLysLysValArgValLysCysAsnThr 20

Qy  61  GATGATACCATCGGGACCTTAAAGAGCTGATTGCAGCC 99
Db  21  AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33

RESULT 8
US-09-331-930A-22
; Sequence 22, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007, 00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11

```

```

; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-331-930A-22

Alignment Scores:
Pred. No.: 1.17e-12 Length: 73
Score: 135.00 Matches: 26
Percent Similarity: 84.8% Conservative: 2
Best Local Similarity: 78.8% Mismatches: 5
Query Match: 71.8% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x US-09-331-930A-22 (1-73)
QY 1 ATGATCGAGTGTCTTTGCAACGACCTCTGGGGAAAAAGTCCTTAATGCAACACG 60
Db 1 MetIleGluIleThrValAsnAspArgLeuGlyLysValArgIleLysCysAsnPro 20
QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99
Db 21 SerAspThrIleGlyAspLeuLysLeuIleAlaA 33

RESULT 9
US-09-331-930A-23
; Sequence 23, Application US/09331930A
; Patent No. 6436670
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL Z.
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 22975-20007.00
; CURRENT APPLICATION NUMBER: US/09/331,930A
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP0117/97
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP0323/97
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Fasiola hepatica
US-09-331-930A-23

Alignment Scores:
Pred. No.: 4.01e-08 Length: 66
Score: 106.00 Matches: 21
Percent Similarity: 84.6% Conservative: 1
Best Local Similarity: 80.8% Mismatches: 4
Query Match: 56.4% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x US-09-331-930A-23 (1-66)
QY 22 GACCGTCTGGGAAAAAGTCCTTAATGCAACGATGATACCATCGGGACCTT 81
Db 1 AspArgLeuGlyLysValArgValLysCysAsnProThrAspLysValGlyAspLeu 20
QY 82 AAGAAGCTGATTGCAGCC 99
Db 21 LysLysLeuIleAlaA 26

RESULT 10
US-09-248-796A-16302
; Sequence 16302, Application US/09248796A
; Patent No. 6747137
```

```

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16302
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16302

Alignment Scores:
Pred. No.: 4.09e-08 Length: 75
Score: 106.00 Matches: 19
Percent Similarity: 80.6% Conservative: 6
Best Local Similarity: 61.3% Mismatches: 6
Query Match: 56.4% Indels: 0
DB: 2 Gaps: 0

US-10-067-832D-13 (1-102) x US-09-248-796A-16302 (1-75)
QY 1 ATGATCGAGTGTCTTTGCAACGACCTCTGGGGAAAAAGTCCTTAATGCAACACG 60
Db 3 MetIleGluIleGlnAlaAenAspArgLeuGlyLysLysIleLysLeuLysCysLeuGlu 22
QY 61 GATGATACCATCGGGGACCTTAAGAGCTGATT 93
Db 23 ThrAspThrIleGlyAspValLysLysIleLeu 33

RESULT 11
US-09-252-991A-29573
; Sequence 29573, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29573
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29573

Alignment Scores:
Pred. No.: 0.971 Length: 92
Score: 59.00 Matches: 10
Percent Similarity: 56.5% Conservative: 3
Best Local Similarity: 43.5% Mismatches: 8
Query Match: 31.4% Indels: 2
DB: 2 Gaps: 1

US-10-067-832D-13 (1-102) x US-09-252-991A-29573 (1-92)
QY 73 CGATGCTATCATCGGTGTCATTTAACGNGGACCTTTTCCCGACGAGCGTCGTTCGAAA 14
Db 58 ArgtrpSerArgProCysAlaValProAlaGlyProLeuAla-----AlaAArgCysArg 75
QY 13 CACCTCGA 5
```

```
Db          76 AspProArg 78
|||||
RESULT 12
US-09-248-796A-17325
; Sequence 17325, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17325
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17325
Alignment Scores:
Pred. No.: 1.88      Length: 644
Score: 58.00      Matches: 12
Percent Similarity: 63.0%      Conservative: 5
Best Local Similarity: 44.4%      Mismatches: 10
Query Match: 30.9%      Indels: 0
DB: 2      Gaps: 0

US-10-067-832D-13 (1-102) x US-09-248-796A-17325 (1-644)
QY 81 AAGTCCCGCATGATATCCGTTGTCATTTAACGNGGACCTTTTCCCGACGGTC 22
|||||
Db 588 LysAlaProAspGlyLeuIlePheArgAlaPheAsnLysAsnTyrSerThrAspThrLeu 607
QY 21 GTTCAACAAACCTCGATCAT 1
|||||
Db 608 ValMetThrSerLeuAspTyr 614

RESULT 13
PCT-US96-01735-1
; Sequence 1, Application PC/TUS9601735
; GENERAL INFORMATION:
; APPLICANT: Marks, Andrew R.
; TITLE OF INVENTION: HUMAN T CELL INOSITOL 1,4,5,-TRISPHOSPHATE
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01735
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/386,039
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kole, Lies B
; REGISTRATION NUMBER: 35,225

REFERENCE/DOCKET NUMBER: A30042 - 165/30555
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2628
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2713 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
PCT-US96-01735-1
Alignment Scores:
Pred. No.: 5.77      Length: 2713
Score: 55.50      Matches: 14
Percent Similarity: 58.8%      Conservative: 6
Best Local Similarity: 41.2%      Mismatches: 11
Query Match: 29.5%      Indels: 3
DB: 4      Gaps: 2

US-10-067-832D-13 (1-102) x PCT-US96-01735-1 (1-2713)
QY 96 TGCATCAGCTTCTTAAGTCCCGCATGATATCATCCGTGTGTCATTTAACGNGGACCTT 37
|||||
Db 1398 CysAsnSerLeuLeuProLeu---AspAspIleValArgValValThrHisGluAspCys 1416
QY 36 TTTCGCCAGACGGTCGTTGCA-----AACAACTCGATCAT 1
|||||
Db 1417 IleProGluValIleAlaTyrIleAsnPheLeuAsnHis 1430

RESULT 14
US-09-385-222A-4
; Sequence 4, Application US/09385222A
; Patent No. 6465211
; GENERAL INFORMATION:
; APPLICANT: RIKEN;
; APPLICANT: Katsuhiko Mikoshiba
; TITLE OF INVENTION: High affinity IP3 binding polypeptide
; FILE REFERENCE: PH-657US
; CURRENT APPLICATION NUMBER: US/09/385,222A
; CURRENT FILING DATE: 1999-08-26
; PRIOR APPLICATION NUMBER: JP98/242207
; PRIOR FILING DATE: 1998-08-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2749
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-385-222A-4
Alignment Scores:
Pred. No.: 5.78      Length: 2749
Score: 55.50      Matches: 14
Percent Similarity: 58.8%      Conservative: 6
Best Local Similarity: 41.2%      Mismatches: 11
Query Match: 29.5%      Indels: 3
DB: 2      Gaps: 2

US-10-067-832D-13 (1-102) x US-09-385-222A-4 (1-2749)
QY 96 TGCATCAGCTTCTTAAGTCCCGCATGATATCATCCGTGTGTCATTTAACGNGGACCTT 37
|||||
Db 1396 CysAsnSerLeuLeuProLeu---AspAspIleValArgValValThrHisGluAspCys 1414
QY 36 TTTCGCCAGACGGTCGTTGCA-----AACAACTCGATCAT 1
```

```

Db      1415 IleProGluValLysIleAlaTyrIleAsnPheLeuAsnHis 1428
          |||:::      :::|||      |||      |||:::|||
RESULT 15
US-09-328-352-6312
; Sequence 6312, Application US/09328352
; Patent No: 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6312
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6312

Alignment Scores:
Pred. No.:      6.29      Length:      456
Score:          54.50      Matches:     11
Percent Similarity: 61.3%      Conservative: 8
Best Local Similarity: 35.5%      Mismatches: 9
Query Match:      29.0%      Indels:    3
DB:                2          Gaps:       1

US-10-067-832D-13 (1-102) x US-09-328-352-6312 (1-456)

QY      84 CTTAAGTCCCGATCGT-----ATCATCGGTGTCATTTAACGGGACCTTTT 34
          ||:::|||||
Db      178 LeuArgLeuProAspGlyGluArgPheMetLeuArgPheAspArgAlaGluLeuAla 197
          ||::: |||
QY      33 CCCAGACGTCGTTCGAAACACCTCGATCAT 1
          ||::: |||
Db      198 ProArgAspIleValAlaArgThrIleAspHis 208
          ||::: |||

```

Search completed: April 7, 2006, 23:49:31  
Job time : 13.8051 secs



GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:48:08 ; Search time 7.41229 Seconds  
 (without alignments)  
 1149.944 Million cell updates/sec

Title: US-10-067-832D-13

Perfect score: 188

Sequence: 1 atgatcaggtgtgttgcaa.....agaagctgattgcagcctaa 102

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp

-Q=/abs/ABSSWEB spool/US10067832/runat 07042006 175617 9835/app\_query.fasta\_1

-DB=Published Applications AA Main -QFMT=faстан -SUFFIX=n2p.rapbm

-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1

-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct

-THR MAX=100 -THR MIN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext

-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p

-USER=US10067832 @CGN 1 1 405 @runat 07042006 175617 9835 -NCPU=6 -ICPU=3

-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMSOOT=120

-WEN TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	163	86.7	33	5	US-10-067-832D-14
2	163	86.7	68	4	US-10-424-599-250542
3	163	86.7	73	5	US-10-067-832D-2
4	163	86.7	73	5	US-10-067-832D-15
5	163	86.7	73	5	US-10-067-832D-16
6	163	86.7	73	5	US-10-067-832D-19
7	163	86.7	73	5	US-10-067-832D-20
8	163	86.7	73	5	US-10-965-898-68
9	163	86.7	115	4	US-10-264-049-2463
10	152	80.9	73	6	US-11-097-143-6522
11	146	77.7	73	4	US-10-424-599-212296

12	146	77.7	73	4	US-10-437-963-147593
13	146	77.7	73	4	US-10-767-701-46049
14	146	77.7	73	4	US-10-425-115-233843
15	146	77.7	73	4	US-10-425-115-233846
16	146	77.7	73	4	US-10-425-115-233848
17	146	77.7	73	4	US-10-425-115-233849
18	146	77.7	73	4	US-10-425-115-233851
19	146	77.7	73	4	US-10-425-115-233853
20	146	77.7	73	4	US-10-425-115-288480
21	146	77.7	73	4	US-10-425-115-288482
22	146	77.7	73	4	US-10-425-115-288483
23	146	77.7	73	4	US-10-425-115-288486
24	140	74.5	73	4	US-10-425-115-288481
25	137	72.9	73	4	US-10-425-115-288479
26	135	71.8	73	5	US-10-067-832D-17
27	134	71.3	89	4	US-10-425-115-355452
28	131.5	69.9	74	4	US-10-424-599-207467
29	130.5	69.4	73	5	US-10-739-930-8461
30	120	63.8	45	4	US-10-425-115-262612
31	118	62.8	71	4	US-10-424-599-147197
32	114	60.6	89	4	US-10-083-357-770
33	114	60.6	320	4	US-10-338-411-11
34	111	59.0	82	4	US-10-389-640-11
35	110	58.5	73	4	US-10-425-115-290033
36	106	56.4	66	5	US-10-067-832D-18
37	97	51.6	55	4	US-10-425-115-270723
38	97	51.6	110	4	US-10-424-599-213543
39	93	49.5	83	4	US-10-437-963-135926
40	93	49.5	104	4	US-10-083-357-1200
41	86	45.7	59	4	US-10-425-115-356308
42	84	44.7	79	4	US-10-425-115-269824
43	84	44.7	178	4	US-10-437-963-149391
44	83	44.1	110	4	US-10-425-114-60599
45					

## ALIGNMENTS

## RESULT 1

US-10-067-832D-14  
 ; Sequence 14, Application US/10067832D  
 ; Publication No. US20050059108A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZIMMET, PAUL ZEV  
 ; APPLICANT: COLLIER, GREGORY  
 ; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
 ; FILE REFERENCE: 229752000701  
 ; CURRENT APPLICATION NUMBER: US/10/067,832D  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: 09/331,930  
 ; PRIOR FILING DATE: 1999-06-30  
 ; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
 ; PRIOR FILING DATE: 1998-10-30  
 ; PRIOR APPLICATION NUMBER: AU PP 0117  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: AU PP 0323  
 ; PRIOR FILING DATE: 1997-11-11  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 14  
 ; LENGTH: 33  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: MOD RES  
 ; LOCATION: (15)  
 ; OTHER INFORMATION: Leu, Pro, His, or Arg  
 ; US-10-067-832D-14

Alignment Scores:  
 Pred. No.: 4.98e-15 Length: 33  
 Score: 163.00 Matches: 33  
 Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 86.7% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-067-832D-14 (1-33)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysVal\*\*VallysCysAsnThr 20

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrileGlyAspLeuLysLysLeuIleAlaAa 33

## RESULT 2

US-10-424-599-250542  
; Sequence 250542, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 250542  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_6826C.1.pep  
US-10-424-599-250542

Alignment Scores:  
Pred. No.: 5.32e-15 Length: 68  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-424-599-250542 (1-68)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrileGlyAspLeuLysLysLeuIleAlaAa 33

## RESULT 3

US-10-067-832D-2  
; Sequence 2, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975200701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323

; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 2  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Psammomys obesus  
US-10-067-832D-2

## Alignment Scores:

Pred. No.: 5.36e-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-067-832D-2 (1-73)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrileGlyAspLeuLysLysLeuIleAlaAa 33

## RESULT 4

US-10-067-832D-15  
; Sequence 15, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 22975200701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 15  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-067-832D-15

## Alignment Scores:

Pred. No.: 5.36e-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-067-832D-15 (1-73)

QY 1 ATGATCGAGGTGTTTTCGACGACCGTCTGGGAAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValCysAsnAspArgLeuGlyLysValArgVallysCysAsnThr 20

QY 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrileGlyAspLeuLysLysLeuIleAlaAa 33

## RESULT 5

Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-067-832D-19 (1-73)

QY 1 ATGATCGAGTGTGTTTGGCAACGACCGTCTGGGGAAAAAGGTCNCNGTTAAATGCAACACG 60  
|||||  
Db 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
|||||

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
|||||

RESULT 7

US-10-067-832D-20  
; Sequence 20, Application US/10067832D  
; Publication No. US20050059108A1  
; GENERAL INFORMATION:  
; APPLICANT: ZIMMET, PAUL ZEV  
; APPLICANT: COLLIER, GREGORY  
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR  
; FILE REFERENCE: 229752000701  
; CURRENT APPLICATION NUMBER: US/10/067,832D  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 09/331,930  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: PCT/AU98/00902  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: AU PP 0117  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: AU PP 0323  
; PRIOR FILING DATE: 1997-11-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 3.2  
; SEQ ID NO 20  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae

US-10-067-832D-20

Alignment Scores:  
Pred. No.: 5,368-15 Length: 73  
Score: 163.00 Matches: 32  
Percent Similarity: 97.0% Conservative: 0  
Best Local Similarity: 97.0% Mismatches: 1  
Query Match: 86.7% Indels: 0  
DB: 5 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-067-832D-20 (1-73)

QY 1 ATGATCGAGTGTGTTTGGCAACGACCGTCTGGGGAAAAAGGTCNCNGTTAAATGCAACACG 60  
|||||  
Db 1 MetIleGluValValCysAsnAspArgLeuGlyLysLysValArgValLysCysAsnThr 20  
|||||

QY 61 GATGATACCATCGGGGACCTTAAGAAGCTGATTGCAGCC 99  
|||||  
Db 21 AspAspThrIleGlyAspLeuLysLysLeuIleAlaAla 33  
|||||

RESULT 8

US-10-965-898-68  
; Sequence 68, Application US/10965898  
; Publication No. US20050084936A1  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; Bandman, Olga  
; Hillman, Jennifer L.  
; Au-Young, Janice  
; Tang, Y. Tom  
; Yue, Henry  
; Shah, Purvi



Qy 1 ATGATCGAGGTGTTTGCACACGCGCTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluIleThrCysAsnAspArgLeuGlyLysValArgValLysCysAsnPro 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAla 33

## RESULT 11

US-10-424-599-212296  
; Sequence 212296, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 212296  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_33729C.1.p  
US-10-424-599-212296

Alignment Scores:  
Pred. No.: 1.72e-12 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-424-599-212296 (1-73)

Qy 1 ATGATCGAGGTGTTTGCACACGCGCTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValLeuAsnAspArgLeuGlyLysValArgValLysCysAsnAsp 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAla 33

## RESULT 12

US-10-437-963-147593  
; Sequence 147593, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 147593  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Oryza sativa

; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48107C.1.p  
US-10-437-963-147593

Alignment Scores:  
Pred. No.: 1.72e-12 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-437-963-147593 (1-73)

Qy 1 ATGATCGAGGTGTTTGCACACGCGCTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValLeuAsnAspArgLeuGlyLysValArgValLysCysAsnGlu 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAla 33

## RESULT 13

US-10-767-701-46049  
; Sequence 46049, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 46049  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C594\_1.p  
US-10-767-701-46049

Alignment Scores:  
Pred. No.: 1.72e-12 Length: 73  
Score: 146.00 Matches: 29  
Percent Similarity: 90.9% Conservative: 1  
Best Local Similarity: 87.9% Mismatches: 3  
Query Match: 77.7% Indels: 0  
DB: 4 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-767-701-46049 (1-73)

Qy 1 ATGATCGAGGTGTTTGCACACGCGCTCTGGGAAAAGGTCNCGTTAAATGCAACACG 60  
Db 1 MetileGluValValLeuAsnAspArgLeuGlyLysValArgValLysCysAsnGlu 20  
Qy 61 GATGATACCATCGGGACCTTAAGAAGCTGATTGCAGCC 99  
Db 21 AspAspThrIleGlyAspLeuLysLeuIleAlaAla 33

## RESULT 14

US-10-425-115-233843  
; Sequence 233843, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 147593  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Oryza sativa

Job time : 38.0614 secs

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 233843
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144858C.1.pep
US-10-425-115-233843
```

```
Alignment Scores:
Pred. No.:      1.72e-12      Length:      73
Score:          146.00      Matches:      29
Percent Similarity: 90.9%      Conservative: 1
Best Local Similarity: 87.9%      Mismatches: 3
Query Match:      77.7%      Indels:      0
DB:              4          Gaps:      0
```

US-10-067-832D-13 (1-102) x US-10-425-115-233843 (1-73)

```
QY      1 ATGATCGAGTTGTTTGCACACCGCTCTGGGAAAAGGTCNCGTTAAATGCACACG 60
      |||||||
Db      1 MetilegluValValleuAsnAspArgLeuGlyLysValArgValLysCysAsnGlu 20
QY      61 GATGATACCATCGGGACCTTAAGACGCTGATTGCAGCC 99
      |||||||
Db      21 AspAspThrileGlyAspLeuLysLeuValAlaAla 33
```

## RESULT 15

```
US-10-425-115-233846
; Sequence 233846, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 233846
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144860C.1.pep
US-10-425-115-233846
```

```
Alignment Scores:
Pred. No.:      1.72e-12      Length:      73
Score:          146.00      Matches:      29
Percent Similarity: 90.9%      Conservative: 1
Best Local Similarity: 87.9%      Mismatches: 3
Query Match:      77.7%      Indels:      0
DB:              4          Gaps:      0
```

US-10-067-832D-13 (1-102) x US-10-425-115-233846 (1-73)

```
QY      1 ATGATCGAGTTGTTTGCACACCGCTCTGGGAAAAGGTCNCGTTAAATGCACACG 60
      |||||||
Db      1 MetilegluValValleuAsnAspArgLeuGlyLysValArgValLysCysAsnGlu 20
QY      61 GATGATACCATCGGGACCTTAAGACGCTGATTGCAGCC 99
      |||||||
Db      21 AspAspThrileGlyAspLeuLysLeuValAlaAla 33
```

Search completed: April 7, 2006, 23:55:18

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 7, 2006, 23:49:48 ; Search time 0.994068 Seconds  
(without alignments)  
640.114 Million cell updates/sec

Title: US-10-067-832D-13

Perfect score: 188

Sequence: 1 atgacgaggtgtttgcaa.....agaagctgattgcgcctaa 102

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 368322

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10067832 @CGN 1 1 34 @runat 07042006 175620 9893/app query.fasta.1  
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs02p  
-USER=US10067832 @CGN 1 1 34 @runat 07042006 175620 9893 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA New:

1: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pep.\*  
2: /SID55/ptodata/1/pubpaa/US06 NEW PUB.pep.\*  
3: /SID55/ptodata/1/pubpaa/US07 NEW PUB.pep.\*  
4: /SID55/ptodata/1/pubpaa/PTC\_NEW PUB.pep.\*  
5: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep.\*  
6: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep.\*  
7: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep.\*  
8: /SID55/ptodata/1/pubpaa/US60\_NEW PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56.5	30.1	258	7	US-11-058-924-6
2	52.5	27.9	213	7	US-11-072-512-2722
3	52	27.7	371	7	US-11-096-568A-5963
4	51.5	27.4	937	7	US-11-098-686-11296
5	51	27.1	37	6	US-10-957-351-107
6	51	27.1	4544	6	US-10-501-035-214
7	51	27.1	4544	7	US-11-076-427A-32
8	50	26.6	319	7	US-11-096-568A-17313
9	50	26.6	346	7	US-11-096-568A-17312

10	50	26.6	375	7	US-11-096-568A-17311	Sequence 17311, A
11	50	26.6	525	6	US-10-467-657-3462	Sequence 5462, Ap
12	49	26.1	3475	7	US-11-087-099-10885	Sequence 10885, A
13	49	26.1	3488	7	US-11-087-099-3005	Sequence 9005, Ap
14	48.5	25.8	1299	6	US-10-821-234-1145	Sequence 1145, Ap
15	48	25.5	47	7	US-11-004-399-3161	Sequence 3161, Ap
16	48	25.5	90	6	US-10-644-807-237	Sequence 237, App
17	48	25.5	261	7	US-11-096-568A-5123	Sequence 5123, Ap
18	48	25.5	359	7	US-11-123-013-6	Sequence 6, Appl
19	47	25.0	114	7	US-11-096-568A-31955	Sequence 31955, A
20	47	25.0	175	7	US-11-096-568A-31954	Sequence 31954, A
21	47	25.0	250	7	US-11-096-568A-30582	Sequence 30582, A
22	47	25.0	258	7	US-11-096-568A-30581	Sequence 30581, A
23	47	25.0	272	7	US-11-096-568A-30580	Sequence 30580, A
24	47	25.0	286	7	US-11-096-568A-31559	Sequence 31559, A
25	47	25.0	294	7	US-11-096-568A-31558	Sequence 31558, A
26	47	25.0	296	6	US-10-467-657-5024	Sequence 5024, Ap
27	47	25.0	308	7	US-11-096-568A-31557	Sequence 31557, A
28	47	25.0	355	7	US-11-087-099-3295	Sequence 3295, Ap
29	47	25.0	584	7	US-11-096-568A-31265	Sequence 31265, A
30	47	25.0	610	7	US-11-096-568A-31264	Sequence 31264, A
31	47	25.0	640	7	US-11-096-568A-31263	Sequence 31263, A
32	47	25.0	1126	7	US-11-075-185-3	Sequence 3, Appl
33	46.5	24.7	274	7	US-11-087-099-6337	Sequence 6337, Ap
34	46.5	24.7	648	7	US-11-087-099-4135	Sequence 4135, Ap
35	46.5	24.7	964	7	US-11-103-957-13	Sequence 13, Appl
36	46.5	24.7	964	7	US-11-018-868-19	Sequence 19, Appl
37	46.5	24.7	1574	6	US-10-055-877-211	Sequence 211, App
38	46	24.5	70	7	US-11-096-568A-1855	Sequence 1855, Ap
39	46	24.5	247	7	US-11-096-568A-29306	Sequence 29306, A
40	46	24.5	256	7	US-11-096-568A-29305	Sequence 29305, A
41	46	24.5	293	7	US-11-096-568A-16425	Sequence 16425, A
42	46	24.5	311	7	US-11-096-568A-29304	Sequence 29304, A
43	46	24.5	314	7	US-11-096-568A-19764	Sequence 19764, A
44	46	24.5	337	7	US-11-096-568A-19763	Sequence 19763, A
45	46	24.5	342	7	US-11-096-568A-19762	Sequence 19762, A

#### ALIGNMENTS

RESULT 1  
US-11-058-924-6  
; Sequence 6, Application US/11058924  
; Publication No. US20060019308A1  
; GENERAL INFORMATION:  
; APPLICANT: Lund, Frances E.  
; APPLICANT: Randall, Troy D.  
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS  
; FILE REFERENCE: 13315/5  
; CURRENT APPLICATION NUMBER: US/11/058, 924  
; CURRENT FILING DATE: 2005-02-15  
; PRIOR APPLICATION NUMBER: 09/982,616  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: 60/241,065  
; PRIOR FILING DATE: 2000-10-17  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 6  
; LENGTH: 258  
; TYPE: PRT  
; ORGANISM: A. californica  
US-11-058-924-6

Alignment Scores:  
Pred No.: 0.867 Length: 258  
Score: 56.50 Matches: 12  
Percent Similarity: 62.5% Conservative: 8  
Best Local Similarity: 37.5% Mismatches: 11  
Query Match: 30.1% Indels: 1  
DB: 7 Gaps: 1

US-10-067-832D-13 (1-102) x US-11-058-924-6 (1-258)

```
QY 4 ATCGAGGTGTTGTCACGACCGTCTGGGAAAAAGGTCCNCGTTAAATGCAACACGGAT 63
      :::::||||:|  ||| |||||:::||||:|  ||| |||||
Db 190 VallysValIleValLeuHisArgLeuGlyGluLysIlelleGluLysCysGlyAlaGly 209
      ::: |||:::|||||:::|  |||
QY 64 GATACCATCGGACCTTAAGAAGCTGATTGCAGCC 99
      ::: |||:::|||||:::|  |||
Db 210 SerLeuLeu--AspLeuGluLysLeuValLysAla 220
      ::: |||:::|||||:::|  |||

RESULT 2
US-11-072-512-2722
; Sequence 2722, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2722
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2722

Alignment Scores:
Pred. No.: 3.66 Length: 213
Score: 52.50 Matches: 12
Percent Similarity: 41.0% Conservative: 4
Best Local Similarity: 30.8% Mismatches: 8
Query Match: 27.9% Indels: 15
DB: 7 Gaps: 2

US-10-067-832D-13 (1-102) x US-11-072-512-2722 (1-213)
QY 76 CCCCAGTGTATCATCCGTG-----TGC 53
      |||:::|  |||
Db 160 ProLeuTrpAsnPheAsnCysGlnPhePheIleLysAspLeuTyrGlnAspValLeuCys 179
      |||:::|  |||
QY 52 ATT-----TAACGNGGACCTTTTCCCGACAGCGTGTGCGAAACAACTCGA 5
      ::: |||:::|||||:::|  |||
Db 180 LeuThrLeuPheAspArgAspGlnPheSerProAspGlyLysCysArgSerAlaArg 198
      ::: |||:::|||||:::|  |||

RESULT 3
US-11-096-568A-5963
; Sequence 5963, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Thery
```

```
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5963
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(371)
; OTHER INFORMATION: Ceres Seq. ID no. 14312794
US-11-096-568A-5963

Alignment Scores:
Pred. No.: 4.65 Length: 371
Score: 52.00 Matches: 8
Percent Similarity: 91.7% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 1
Query Match: 27.7% Indels: 0
DB: 7 Gaps: 0

US-10-067-832D-13 (1-102) x US-11-096-568A-5963 (1-371)
QY 99 GCGTCAATCAGCTTCTTAAGGTCCCGCATGTATC 64
      ||| |||||  |||:::|  ||| |||||:::|
Db 47 GlyCysAsnSerLeuLeuArgValProGluGlyLeu 58
      ||| |||||  |||:::|  ||| |||||:::|

RESULT 4
US-11-098-686-11296
; Sequence 11296, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11296
; LENGTH: 937
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11296

Alignment Scores:
Pred. No.: 6.12 Length: 937
Score: 51.50 Matches: 13
Percent Similarity: 57.1% Conservative: 3
Best Local Similarity: 46.4% Mismatches: 9
Query Match: 27.4% Indels: 3
DB: 7 Gaps: 1

US-10-067-832D-13 (1-102) x US-11-098-686-11296 (1-937)
QY 86 TTCTTAAGTCCCGCATGTATCATCCGTGTGCAATTAAACGNGGACCTTTTCCCGCAGA 27
      |||:::|  |||:::|  |||:::|  |||:::|
Db 211 PheValArgPheProIleHisAspLysArgLeuHis-----ThrIlePheProGln 227
      |||:::|  |||:::|  |||:::|  |||:::|

QY 26 CGGTCTGTGCAACAACTCGATC 3
      ||| |||||
Db 228 AlaAspLeuThrThrSerIle 235
      ||| |||||

RESULT 5
US-10-957-351-107
; Sequence 107, Application US/10957351
; Publication No. US20060008844A1
```



GENERAL INFORMATION:  
; APPLICANT: Stemmer, Willem P. C.  
; APPLICANT: Perleth, D. Victor  
; APPLICANT: Satyal, Sanjeev  
; APPLICANT: Avidia Research Institute  
; TITLE OF INVENTION: C-Met Kinase Binding Proteins  
; FILE REFERENCE: 022013-001400US  
; CURRENT APPLICATION NUMBER: US/10/957,351  
; CURRENT FILING DATE: 2004-09-30  
; PRIOR APPLICATION NUMBER: US 10/871,602  
; PRIOR FILING DATE: 2004-06-17  
; NUMBER OF SEQ ID NOS: 471  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 107  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: human LDL-receptor class A domain  
US-10-957-351-107

Alignment Scores:  
Pred. No.: 5.32 Length: 37  
Score: 51.00 Matches: 10  
Percent Similarity: 58.3% Conservative: 4  
Best Local Similarity: 41.7% Mismatches: 10  
Query Match: 27.1% Indels: 0  
DB: 6 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-957-351-107 (1-37)

Qy 7 GAGTTGTTTCAACGACCGTCTGGGAAAGTCNCGTAAATGCAACACGGATGAT 66

Db 5 GluPheGlnCysAsnAsnThrLeuCysLysProLeuAlaTrpLysCysAspGlyGluAep 24

Qy 67 ACCATCGGGGAC 78

Db 25 AspCysGlyAep 28

RESULT 6

US-10-501-035-214  
; Sequence 214, Application US/10501035  
; Publication No. US20060046249A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING  
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE  
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS  
; FILE REFERENCE: D0185 PCT  
; CURRENT APPLICATION NUMBER: US/10/501,035  
; CURRENT FILING DATE: 2004-07-09  
; PRIOR APPLICATION NUMBER: US 60/350,061  
; PRIOR FILING DATE: 2002-01-18  
; NUMBER OF SEQ ID NOS: 795  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 214  
; LENGTH: 4544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-501-035-214

Alignment Scores:  
Pred. No.: 8.59 Length: 4544  
Score: 51.00 Matches: 10  
Percent Similarity: 58.3% Conservative: 4  
Best Local Similarity: 41.7% Mismatches: 10  
Query Match: 27.1% Indels: 0  
DB: 6 Gaps: 0

US-10-067-832D-13 (1-102) x US-10-501-035-214 (1-4544)

Qy 7 GAGTTGTTTCAACGACCGTCTGGGAAAGTCNCGTAAATGCAACACGGATGAT 66

Db 5 GluPheGlnCysAsnAsnThrLeuCysLysProLeuAlaTrpLysCysAspGlyGluAep 24

Db 3658 GluPheGlnCysAsnAsnThrLeuCysLysProLeuAlaTrpLysCysAspGlyGluAep 3677  
Qy 67 ACCATCGGGGAC 78  
Db 3678 AspCysGlyAep 3681

RESULT 7

US-11-076-427A-32  
; Sequence 32, Application US/11076427A  
; Publication No. US20060025338A1  
; GENERAL INFORMATION:  
; APPLICANT: Alicato, et al.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF LYMPHATIC AND VENOUS  
; TITLE OF INVENTION: VESSEL ARTERIALIZATION  
; FILE REFERENCE: 28967/40008A  
; CURRENT APPLICATION NUMBER: US/11/076,427A  
; CURRENT FILING DATE: 2005-03-08  
; PRIOR APPLICATION NUMBER: US 60/551,581  
; PRIOR FILING DATE: 2004-03-08  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 32  
; LENGTH: 4544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-076-427A-32

Alignment Scores:  
Pred. No.: 8.59 Length: 4544  
Score: 51.00 Matches: 10  
Percent Similarity: 58.3% Conservative: 4  
Best Local Similarity: 41.7% Mismatches: 10  
Query Match: 27.1% Indels: 0  
DB: 7 Gaps: 0

US-10-067-832D-13 (1-102) x US-11-076-427A-32 (1-4544)

Qy 7 GAGTTGTTTCAACGACCGTCTGGGAAAGTCNCGTAAATGCAACACGGATGAT 66

Db 3658 GluPheGlnCysAsnAsnThrLeuCysLysProLeuAlaTrpLysCysAspGlyGluAep 3677

Qy 67 ACCATCGGGGAC 78

Db 3678 AspCysGlyAep 3681

RESULT 8

US-11-096-568A-17313  
; Sequence 17313, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: Therby  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 17313  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(319)  
; OTHER INFORMATION: Ceres Seq. ID no. 12357282  
US-11-096-568A-17313

Alignment Scores:  
Pred. No.: 9.5 Length: 319  
Score: 50.00 Matches: 12  
Percent Similarity: 63.3% Conservative: 7  
Best Local Similarity: 40.0% Mismatches: 7  
Query Match: 26.6% Indels: 4



```
; LENGTH: 3475
; TYPE: PRT
; ORGANISM: Nodularia spumigena
US-11-087-099-10885

Alignment Scores:
Pred. No.:      17.4          Length:    3475
Score:         49.00        Matches:     11
Percent Similarity: 62.1%   Conservative: 7
Best Local Similarity: 37.9% Mismatches:   10
Query Match:    26.1%       Indels:       1
DB:             7           Gaps:         0

US-10-067-832D-13 (1-102) x US-11-087-099-10885 (1-3475)

QY 98 GTTGC AATCAGCTTC TTA AGGTCCCGCATG CATCCGGTG TCATTAAACGGGACC 39
||| :::: :: :||| :::: ||| :::: |||
Db 231 AlaGIyValAsnLeuIleAsnProAlaIleThrIleAsnLeuSerGlnSer-GlyMe 250

QY 38 TTTTTC CCAAGCGGTC GTTGC AAA 14
||||| | | | | | | | | | | | | |
Db 250 tMetSerProAspGlyArgCysLys 258

RESULT 13
US-11-087-099-9005
; Sequence 9005, Application US/11087099
; Publication NO. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450) B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9005
; LENGTH: 3488
; TYPE: PRT
; ORGANISM: Planctothrix agardhi
US-11-087-099-9005

Alignment Scores:
Pred. No.:      17.4          Length:    3488
Score:         49.00        Matches:     11
Percent Similarity: 62.1%   Conservative: 7
Best Local Similarity: 37.9% Mismatches:   10
Query Match:    26.1%       Indels:       1
DB:             7           Gaps:         0

US-10-067-832D-13 (1-102) x US-11-087-099-9005 (1-3488)

QY 98 GTTGC AATCAGCTTC TTA AGGTCCCGCATG CATCCGGTG TCATTAAACGGGACC 39
||| :::: :: :||| :::: ||| :::: |||
Db 231 AlaGIyValAsnLeuIleAsnProAlaIleThrIleAsnLeuSerGlnSer-GlyMe 250

QY 38 TTTTTC CCAAGCGGTC GTTGC AAA 14
||||| | | | | | | | | | | | | |
Db 250 tLeuSerProAspGlyArgCysLys 258

RESULT 14
US-10-821-234-1145
; Sequence 1145, Application US/10821234
; Publication NO. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
```

**THIS PAGE BLANK (USPTO)**